

**INTERRATER RELIABILITY OF THE COMPREHENSIVE
ASSESSMENT OF PSYCHOPATHIC PERSONALITY
DISORDER AMONG A SAMPLE OF INCARCERATED
SERIOUS AND VIOLENT YOUNG OFFENDERS**

by

Amanda McCormick
Bachelor of Arts, Simon Fraser University, 2004

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

In the
School of Criminology

© Amanda McCormick 2007

SIMON FRASER UNIVERSITY

2007

All rights reserved. This work may not be
reproduced in whole or in part, by photocopy
or other means, without permission of the author.

APPROVAL

Name: Amanda McCormick
Degree: Master of Arts
Title of Thesis: Interrater Reliability of the Comprehensive Assessment of Psychopathic Personality Disorder among a Sample of Incarcerated Serious and Violent Young Offenders

Examining Committee:

Chair: Brian Burtch, PhD
Professor of Criminology

Raymond R. Corrado, PhD
Senior Supervisor
Professor, School of Criminology

Patrick Lussier, PhD
Supervisor
Assistant Professor, School of Criminology

Stephen D. Hart, PhD
Professor, Department of Psychology, Simon Fraser University

Date Defended/Approved:

July 30, 2007



SIMON FRASER UNIVERSITY
LIBRARY

Declaration of Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the "Institutional Repository" link of the SFU Library website <www.lib.sfu.ca> at: <<http://ir.lib.sfu.ca/handle/1892/112>>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author's written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada



SIMON FRASER UNIVERSITY
THINKING OF THE WORLD

STATEMENT OF ETHICS APPROVAL

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

(a) Human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

(b) Advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

(c) as a co-investigator, in a research project approved in advance,

or

(d) as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Bennett Library
Simon Fraser University
Burnaby, BC, Canada

ABSTRACT

Over the past several decades, the Psychopathy Checklist (PCL) has demonstrated utility in predicting general and violent criminality. However, the PCL has been criticized for over-reliance on potentially tautological characteristics such as criminal behaviour to assess its validity, as well as the absence of conceptually related characteristics such as a lack of anxiety. Such criticisms are central to any reconceptualizations of psychopathic personality disorder (PPD). The Comprehensive Assessment of Psychopathic Personality Disorder (CAPP) presents a hierarchical conceptual model of psychopathy with six dimensions: attachment, behavioural, cognitive, dominance, emotional, and self-styles of functioning. This thesis analyzes the interrater reliability of the CAPP with a sample of 30 incarcerated youth. The results indicate that the CAPP total scores have excellent interrater reliability while the domain scores have good to excellent interrater reliability. Utilizing this data set, future research will include the validation of the CAPP as a more comprehensive measure of PPD.

Keywords: Young offenders; psychopathy; personality disorder; CAPP.

ACKNOWLEDGMENTS

I would like to thank several individuals who, without their support, this thesis would not have been possible.

To my parents, Nigel and Vivien Watkinson, my sister, Jaclyn Watkinson, and my husband, Adam McCormick, thank you for your never-ending patience, support, and interest in my work.

To my supervisors, Raymond Corrado and Patrick Lussier, I would like to thank you for your guidance and willingness to share your knowledge with me. Under your mentorship, I have learned many things that I would not have thought possible, and I thank you for offering me these opportunities. In addition, I would like to thank Ray Corrado and Irwin Cohen for the opportunity to manage their youth project, which provided me with the necessary information to conduct my thesis analyses. You have both provided me with so many astounding opportunities, for which I will be forever grateful.

To the authors of the CAPP, David Cooke, Stephen Hart, and Caroline Logan, I would like to express my appreciation towards you for providing us with this unpublished instrument, enabling us to provide the initial results with respect to its use. In particular, I would like to thank Dr. Hart for the time spent in training myself and others in the use of this instrument.

Lastly, I would like to acknowledge the contribution of the Social Sciences and Humanities Research Council of Canada, who provided me with a Master's Scholarship with which to conduct my research.

TABLE OF CONTENTS

Approval	ii
Abstract	iii
Acknowledgments	iv
TABLE OF CONTENTS	v
List of Figures and Tables	vi
The Construct of Psychopathy	1
Psychopathy as a Criminological Theory	2
The Psychopathy Checklist	5
Psychopathy and Young Offenders	9
Validity of the Psychopathy Construct in Youth	13
Subtypes of Psychopathy	23
Psychopathy: A Categorical or Dimensional Construct?	27
The Comprehensive Assessment of Psychopathic Personality Disorder	31
The CAPP Model of Psychopathy	35
Validity and Reliability	46
Validity	46
Reliability	49
Intraclass Correlation Coefficients	53
Method	56
Participants	56
Procedure	62
Instrument	64
Statistical Analysis	67
Results	71
Description of CAPP Scores	71
Moderator Variables	74
Interrater Reliability	77
Correlations between Dimensions of the CAPP	82
Discussion	85
Methodological issues and limitations	91
Future Research	94
Conclusion	97
References	100

LIST OF FIGURES AND TABLES

Figure 1: Development in the Models of Psychopathy.....	37
Table 1: Cleckley's 16 Item Model of Psychopathy (adapted from Cleckley, 1964, pp: 362-363).....	36
Table 2: Hare's 20-item and 13-item Models of Psychopathy (adapted from Hare & Neumann, 2006 and Corrado, Vincent, Hart, & Cohen, 2004)	39
Table 3: Cooke, Hart, and Logan (unpublished manuscript) 6-Dimensional Model of Psychopathy (the CAPP Domains, Symptoms, and Illustrative Indicators).....	40
Table 4: Sample ICCs for Total Score Interrater Reliability of PCL Measures	54
Table 5: Demographic Information of Sample	59
Table 6: Past Charges for Sample of Serious and Violent Incarcerated Youth (n = 28)	61
Table 7: Most Serious Charge Resulting in Current Incarceration for Sample	62
Table 8: Guidelines for Assessing Interrater Reliability using Intraclass Correlation Coefficients.....	67
Table 9: Descriptives of CAPP Domains and Total Scores.....	71
Table 10: Internal Consistency Reliability of CAPP Total Score from Domain Scores.....	72
Table 11: Internal Consistency Reliabilities of Domain Scores from Items	73
Table 12: Moderator Variable Effects on the CAPP Total Scores	76
Table 13: Moderator Variable Effects on CAPP Domain Scores.....	77
Table 14: Total and Domain Intraclass Correlation Coefficients for all Rater Pairs.....	78
Table 15: Item Intraclass Correlation Coefficients for all Rater Pairs.....	79
Table 16: Total, Domain, and Item Intraclass Correlation Coefficients for each set of Rater Pairs.....	80
Table 17: Item Intraclass Correlation Coefficients for each set of Rater Pairs	81
Table 18: Correlations between Domain and Total CAPP Scores	84

THE CONSTRUCT OF PSYCHOPATHY

Personality refers to one's way of thinking, feeling, and acting towards others (Lynam & Derefinko, 2006). Essentially, personality represents a pattern of these three dimensions. Psychopathy is a personality disorder where an individual typically displays dysfunctional affective, interpersonal, and behavioural traits when relating to others. Psychopaths are depicted as cold and callous individuals who experience no genuine empathy and remorse despite their harmful actions. The behaviour of psychopaths is routinely impulsive and irresponsible, frequently resulting in contact with the criminal justice and/or mental health systems. Psychopaths are manipulative, lying easily and often. Moreover, they are superficial, glib, and egocentric, giving the impression that "something is off" about them (Hare, 1998; Hare, 2003; Cooke & Michie, 2001). Psychopathic Personality Disorder (PPD) is primarily manifested during interpersonal interactions with others. Cleckley summarized this interaction when he wrote that "the features that are most important in the behavior of the psychopath do not adequately emerge when their behavior is relatively isolated. The qualities [traits] of the psychopath become manifest only when he is connected into the full circuits of social life" (Cleckley, 1964: 40).

The recognized "gold standard" in the assessment of PPD is the Psychopathy Checklist (PCL) developed by Dr. Robert Hare. The PCL has been revised several times (PCL-R; Hare, 1991, 2003), and has since been further developed into a youth version (PCL:YV; Forth, Kosson, & Hare, 2003) and a screening version for use within non-forensic samples (PCL:SV; Hart, Cox, & Hare, 1995). While the PCL scales have produced a substantial body of empirical research contributing towards our understanding

of psychopathy, they are also limited, due to an over-reliance on criminal behaviours and a failure to assess additional clinically relevant items. In an attempt to further the understanding of the content of psychopathic personality disorder (PPD), a new hierarchical measure of psychopathy, the Comprehensive Assessment of Psychopathy Personality (CAPP) Disorder, was recently developed by Drs. David Cooke, Stephen Hart, and Carolyn Logan. In the current study, the content of this new measure will be described, and the initial results regarding the interrater reliability of the CAPP will be presented using a sample of incarcerated serious and violent young offenders.

Psychopathy as a Criminological Theory

Many proposed theoretical explanations of criminal behaviour rely upon biological factors, sociological factors, and/or environmental factors to predict criminality. One such theory refers to PPD as an explanatory factor of criminal behaviour. For at least a small proportion of those engaging in criminal behaviour, their persistent criminality is theoretically due to the presence of a set of dysfunctional traits that limit their appreciation of the thoughts, concerns, and feelings of others. The defining traits of psychopathy, such as a lack of empathy, lack of remorse, superficiality, and egocentricity, theoretically result in behaviour that is reckless and that lacks consideration for the physical, emotional, or psychological impact it will have on others.

The traits of the psychopath were captured in the manuscripts of Hervey Cleckley. Cleckley's clarification of the psychopath referred to 16 descriptive traits, including superficial charm and good intelligence, untruthfulness and insincerity, pathologic egocentricity and incapacity for love, general poverty in major affective reactions, specific loss of insight, and unresponsiveness in general interpersonal relations (Cleckley,

1964). It is arguably the presence of these traits, among others, that predispose the psychopathic individual to engaging in behaviours that are physically, emotionally, and psychologically harmful to others. Incapacity to feel emotion leads to a subsequent lack of remorse and lack of consideration for the effects that one's behaviour will have on others. The traits of the psychopath are such that they are unable to comprehend the effects of their behaviours; as such, psychopaths have long been associated with both general and violent criminality.

The origins of such traits are unclear at this time (Ogloff, 2006). Among the possible sources of psychopathy include improper socialization, environmental influences, biological dysfunctions, and genetic influences. In effect, researchers have found that the emotional deficits inherent to psychopathy are due to genetic makeup (e.g. Blonigen, Carlson, Krueger, & Patrick, 2003), brain functioning (e.g. Blair, Newman, Mitchell, Richell, Leonard, Morton, & Blair, 2006), and/or childhood traumatic experiences (e.g. Lykken, 1995; Porter, 1996). Alternative theories suggest that psychopathic traits such as poor avoidance learning may be adopted through social learning processes (Lykken, 1957). Farrington (2005) discusses how attachment theories (cold and distant parents), social learning theories (inconsistent parental discipline), and cognitive theories (thought and decision making processes) have all been variously proposed as sources for the origin of psychopathy.

Recent research has utilized twin studies to explore the relative roles of genetics and the environment in the development of psychopathy. Viding, Blair, Moffitt, and Plomin (2005) analyzed results for over 3,000 twins participating in the Twins Early Development Study who provided scores on the callous-unemotional (CU) and antisocial

behaviour (AB) scales. Using teacher-ratings, they separated their sample into a group in which at least one twin exhibited high levels of CU (defined as 1.31 standard deviations above the mean), and a group in which at least one twin exhibited high levels of AB (defined as 1.28 standard deviations above the mean). The AB group was also further divided into two groups of high and low CU traits. The researchers subsequently compared monozygotic (identical) and dizygotic (fraternal) twin pairs to examine the extent to which these traits could be attributed to genetic and/or environmental traits. The results of the study suggested that there is a strong heritability component for CU traits, while AB traits are largely influenced by environmental factors. Specifically, there was a strong genetic influence in the AB group where at least one twin exhibited a high degree of CU, whereas environmental factors appeared to account for the AB in the group characterized by low CU traits. Viding et al. (2005) suggest that future research should continue to explore the role of genetics, in addition to neuroscience, as current research suggests that neuropsychological evidence provides support for the presence of brain dysfunction that may be related to genetic vulnerabilities.

Research has additionally focused on the potential for brain dysfunction that may underlie psychopathic traits. Studies with psychopathic individuals have shown that they tend to process information and emotions differently than do those without high levels of psychopathic traits. The prefrontal cortex, in particular, has been associated with antisocial and violent behaviours. In reviewing the research on brain imaging and psychopathy, Raine and Yang (2006) identify that structures in the brain including the corpus callosum, posterior hippocampus, anterior hippocampus, and prefrontal grey volume may all play a role in psychopathy. Similarly, Blair (2006) focuses on the

dysfunctional role of the amygdala, which, in non-psychopathic individuals, is typically associated with the processing of emotions. In effect, studies such as these suggest that psychopaths tend to focus their attention and process information differently than non-psychopaths. However, research in this area of the field is relatively limited (e.g. Raine & Yang, 2006), and there is much left to discover regarding the role of the brain and cognition with respect to psychopathy.

Despite the lack of a theoretical consensus on the origins of PPD, what is consistently clear from the literature is that there does appear to be a subgroup of individuals who differ meaningfully from others in identifiable ways (Vien & Beech, 2006). As such, measures have been developed that operationalize the construct of psychopathy and allow for its assessment using standardized techniques.

The Psychopathy Checklist

The Psychopathy Checklist-Revised (PCL-R) is a 20-item measure of psychopathy that relies heavily on Cleckley's conceptualization of psychopathy. Factor analyses of data collected with the PCL-R have consistently identified two distinctive but highly correlated (approximately $r = .50$) factors – Factor 1, Affective/Interpersonal, which represents the “core” psychopathic traits of callousness, lack of empathy, superficial charm, and egocentricity, and, Factor 2, Socially Deviant Lifestyle, which represents the behavioural dimension, characterized by impulsive and irresponsible behaviours (Hare, 1991).

Recent theorizing has led to speculations that a two-factor model of psychopathy is insufficient. Subsequent research has suggested that the inclusion of affective and interpersonal traits into a single factor inappropriately limits the conceptualization of

psychopathic personality disorder. Cooke and Michie (2001) provided evidence in support of the separation of psychopathy into a three-factor structure consisting of Arrogant and Deceitful Interpersonal Style, Deficient Affective Experience, and Impulsive and Irresponsible Behavioral Style. In doing so, Cooke and Michie (2001) not only re-organized the conceptualization of psychopathy from two into three main factors, but they also removed antisocial behaviours that they argued were irrelevant to the construct. In effect, of the seven items on the PCL-R removed by Cooke and Michie (2001), six were related in some manner to antisocial behaviour.

Cooke and Michie (2001) supported the removal of these traits by referring to the distinction between basic tendencies (core personality traits) and characteristic adaptations (behavioural manifestations that represent the combination of personality traits and sociocultural influences), implying that criminality may be a consequence of such personality traits, rather than a core feature of psychopathy. Cleckley's original conceptualization of psychopathic individuals did not identify criminality or violence as a defining feature. In contrast, the defining features of psychopathy involve a constellation of dysfunctional personality traits that often, although not always, culminate in criminal and other antisocial behaviours.

In effect, rather than antisocial behaviours being a necessary condition of psychopathy, the presence of psychopathic traits may instead increase the risk of subsequent antisocial behaviours, resulting in the common affiliation between psychopathy and crime. Therefore, the defining features of psychopathy are those related to dysfunctional personality traits, and researchers such as Cooke (Cooke, Michie, Hart, & Clark, 2004) have argued that the antisocial behaviours are *consequences* of these

traits, rather than *symptoms* of the disorder itself. Essentially, personality traits such as impulsiveness result in tendencies to behave in certain ways, for instance, by breaking the rules (Farrington, 2005). Therefore, these underlying personality traits shape expressed behaviours. Furthermore, these expressed behaviours (such as promiscuity) may not only be the result of a single underlying personality trait (such as impulsivity), but also the result of a combination of particular personality traits (such as impulsivity, lack of empathy, and irresponsibility) (Lilienfeld, 1994; Lilienfeld, 1998). As such, according to Cooke et al. (2004), the presence of antisocial behaviours should not be a necessary component in identifying the presence of PPD.

Others have argued that the inclusion of antisocial traits related to criminality results in a tautological measurement of future criminality (Farrington, 2005). Past behaviour is one of the strongest predictors of future behaviour; therefore, in predicting future criminality, clinicians and researchers often refer to past criminality. However, when trying to *explain* the occurrence of criminality with past criminality, the relationship becomes tautological, as one is trying to explain behaviour using the behaviour itself as part of the definition. Furthermore, much of the research in the field of psychopathy has been conducted with incarcerated populations, and so it is highly possible that the association between psychopathy and criminality and/or violence has been overstated.

Given the role that past behaviour plays in predicting future behaviours, others have proposed that there is value to the inclusion of antisocial behavioural traits in psychopathy measurements. Salekin, Brannen, Zalot, Leistico, and Neumann (2006) argue that while the elimination of these traits may clarify the concept of PPD, it limits

the ability of the disorder to predict important criminal justice outcomes such as violence and other criminal offending. Research has shown that it is the behavioural factor of psychopathy that is strongly linked to general and violent recidivism, and that this is likely due to the inclusion of antisocial behaviours in the assessment of psychopathy (e.g. Corrado, Vincent, Hart, & Cohen, 2004). However, as discussed above, if psychopathy is to be used as a *theory* of delinquency, then it should not include antisocial behaviours as a criteria or marker (Farrington, 2005) as this results in a tautological measurement. In effect, while the inclusion of the antisocial behaviours, may improve the ability to *predict* future criminality, it will limit the ability of PPD to *explain* criminality.

Since its initial development, the PCL has been revised several times and has been transformed into a screening version for use in non-incarcerated populations (PCL-SV), as well as into a youth version (PCL:YV). Although based on the 20 items originally included in the PCL-R, the PCL:YV has modified several of the items to be more relevant to youth. For instance, the item “many marital relationships” has been modified to reflect impersonal or unstable sexual relationships, while the criteria for “criminal versatility” has been reduced to four types of criminal behaviours to score a two (definitely present), three types to score a one (possibly present), and less than three to score a zero (not present).

Similar to the ongoing debate in the adult literature, there is also debate with respect to the appropriate factor structure of psychopathy in youth. Research with adolescent samples has supported both the traditional three-factor structure of psychopathy, as well as the two-factor, four-facet structure recently proposed by Hare. Using the PCL:YV with a sample of 130 child and adolescent offenders in Miami,

Salekin et al. (2006) identified support for a modified three-factor model characterized by interpersonal (glib/superficial charm, grandiosity, pathological lying, manipulation), affective (shallow affect, lack of empathy, lack of remorse, failure to accept responsibility), and behavioural (stimulation seeking, impulsivity, irresponsibility, parasitic lifestyle, lacking goals) factors. The researchers also identified support for a four-factor model in which persistent and varied rule breaking (early behavioural problems, serious criminal behaviour, serious violations of conditional release, criminal versatility) characterized the fourth, and weakest, factor.

Psychopathy and Young Offenders

The lack of success in treating psychopathy in adulthood has led to the interest in identifying psychopathic traits in children and adolescents, at a time when such traits may still be malleable (Skeem & Cauffman, 2003). Research with adolescent samples suggests that the prevalence rates of psychopathic traits are relatively similar to adult populations. Vasey, Kotov, Frick, and Loney (2005) identify that approximately 20% to 25% of children and youth with severe conduct problems display a high rate of psychopathic traits. Similarly, research using the PCL:YV with incarcerated male youth populations has indicated that approximately 25% of the population can be identified as having a high number of psychopathic traits (Gretton, Hare, & Catchpole, 2004).

Given that research has identified the prevalence of psychopathic traits in incarcerated samples of youth, it appears that psychopathic traits appear as early as adolescence, possibly even as early as childhood. Farrington (2005) refers to two longitudinal studies conducted in London and in Pittsburgh, which have both suggested that personality and corresponding behaviours are relatively consistent from childhood to

adulthood. Similarly, Rutter (2005) notes that conceptualizations of personality disorders in general all assume a childhood origin. In addition, Cleckley (1964) also presumed a childhood or adolescent origin.

Several studies have provided evidence that psychopathic traits predict serious criminality and violence in young offenders. For instance, Corrado, Vincent, Hart, and Cohen (2004) compared the rates of criminal recidivism between youth with low and high scores on the PCL:YV, and concluded that youth with high scores on the PCL:YV recidivated violently more often, and also reoffended both violently and non-violently significantly sooner following release from custody.

These results support an earlier analysis in which Vincent, Vitacco, Grisso, and Corrado (2003) cluster analyzed incarcerated youth based on their scores on the three factors of the PCL:YV. Varying combinations of scores on the three factors of affective, interpersonal, and behavioural traits yielded four clusters of youth: the psychopathic, impulsive, callous-deceitful, and low-traits groups. The psychopathic cluster scored above the mean on all three of the factors, while the low-traits group scored well below the mean on all factors. The impulsive clusters had high scores on the behavioural factor but not on the affective or interpersonal factors, while the callous-deceitful cluster had high scores on the affective and interpersonal factors, but not on the behavioural factor. Subsequent prospective analyses indicated that half of the psychopathic cluster, as compared to around one-quarter of the other three clusters, reoffended violently within an average of four months following release from custody, thus showing support for the predictive validity of psychopathic traits in young offenders (Vincent et al., 2003).

Gretton, Hare, and Catchpole (2004) rated 157 incarcerated male offenders who had received a court-ordered psychological and psychiatric assessment on the PCL:YV. Unfortunately, the PCL:YV scores for this sample were based solely on a review of file information, as opposed to being supplemented by an interview, which may lead to an underestimation of the interpersonal and affective traits of psychopathy (Gretton, Hare, & Catchpole, 2004). Nevertheless, the researchers found that the PCL:YV scores successfully predicted re-offending over a 10 year period. There was particular support for the prediction of violent offences using PCL:YV total scores; violent offending was significantly predicted with respect to increased frequency and shorter time to re-offending. The total scores were unable to significantly predict non-violent reoffending or sexual reoffending, which the authors argue is due to the relatively high base rate (95%) of the former and the low base rate (11%) of the latter.

Gretton and colleagues (2004) also examined the ability of each of the two factors to predict long-term offending. Factor 1 (interpersonal/affective traits) significantly predicted the frequency of violent offending as well as a shorter time to act violently, while Factor 2 (behavioural traits) significantly predicted nonviolent and violent offences, both with respect to frequency of offending and time to re-offend. Factor 2 scores were found to be more powerful in predicting violent offending, which led the researchers to conclude that there was a clearer association between violent outcomes and the behavioural traits of psychopathy. However, the Factor 2 scores included antisocial behavioural traits that others have argued lead to a tautological measure of offending. It is important to note, however, that these relationships held constant even after controlling

for important effects such as age at first offence, number of previous offences, and number of symptoms reflective of conduct disorder (Gretton, Hare, & Catchpole, 2004).

A recent meta-analysis with studies examining the predictive validity of the PCL scales among young offenders upheld the association between psychopathy and both general and violent recidivism. Edens, Campbell, and Weir (2006) meta-analyzed 21 independent studies conducted between 1990 and 2005 in the United States, Canada, and Sweden, all of which used a PCL scale (e.g. PCL-R adopted for youth, PCL-YV) in some form of predictive capacity. The majority of these studies (15) were conducted with solely male samples, while one was conducted with only females, and the remaining five were conducted with both male and female youth. In their analysis of these 2,867 cases, the researchers identified a moderate but significant relationship between psychopathy and general recidivism ($r_w = .24$) and between psychopathy and violent recidivism ($r_w = .25$). The same relationship was not found for sexual recidivism, likely due to the low base rate of sexual offending (Edens, Campbell, & Weir, 2006). Edens and his colleagues further explored the particular relationship between psychopathy and each individual gender, and found that while psychopathy was significantly related to both general ($r = .25$) and violent ($r = .26$) recidivism for males, the same could not be said for females ($r = .13$ and $r = .10$, respectively). Therefore, Edens et al. (2006) expressed caution regarding the applicability of measures such as the PCL to young female offenders. The researchers also explored the potential mediating effect of ethnicity, and found that while there was no difference among the rates of general recidivism, the relationship between psychopathy and violent recidivism became weaker as the proportion of non-White youth increased.

Edens et al. (2006) also computed effect sizes based on the two-factor structure of psychopathy. While both Factor 1 (Interpersonal/Affective) and Factor 2 (Social Deviance) scores were significantly related to general and violent recidivism, they found a greater effect for the Social Deviance scores. Specifically, Factor 1 scores had a statistically significant positive relationship with general ($r = .18$) and violent ($r = .19$) recidivism, as did Factor 2 scores ($r = .29$ and $r = .26$ respectively). Such results are not surprising, given that others have similarly found stronger effects for the Social Deviance factor of psychopathy. This relationship is likely due to the inclusion of criminal behaviour variables among the Factor 2 items, resulting in what some argue to be a tautology of measurement. Both were negligibly related to sexual recidivism ($r = .03$ and $r = .08$ for Factors 1 and 2 respectively).

Validity of the Psychopathy Construct in Youth

Despite empirical support for the utility of predicting recidivism with the PCL:YV, the validity of applying this construct to youth has been criticized (e.g. Seagrave & Grisso, 2002; Hart, Watt, & Vincent, 2002). Concerns exist not only regarding the potential for misuse in youth forensic settings, such as justification for denial to treatment (Skeem & Cauffman, 2003), but also regarding the actual applicability of the psychopathic construct to youth given that adolescence is characterized by change and development. Traits such as impulsiveness, egocentricity, and manipulateness are common to youth who are undergoing a period of development characterized by identity exploration (Seagrave & Grisso, 2002). In other words, psychopathic traits such as impulsiveness, irresponsible behaviour, and egocentricity are all part of normal

adolescent development; therefore, assessing PPD through the rating of these “normal” traits would lead to an overestimation of psychopathy in youth samples (Salekin, 2006).

Skeem and Cauffman (2003) caution the application of the construct to youth, given that they are undergoing a period of much change and development, in particular with respect to their growing maturity and increasing capacity to consider the long-term effects of their behaviours. Edens, Skeem, Cruise, and Cauffman (2001) similarly express several validity and policy concerns when discussing the capability of assessing psychopathic traits in youth. They too argue that adolescence is a period of substantial development, when youth are engaged in exploring their identity and formulating self-perception. They criticize the application of several PCL-R items to youth, particularly those found in the Socially Deviant Lifestyle factor. Edens and colleagues (2001) contend that, in applying the PCL-R with youth samples, some researchers have considered that certain items such as “many short-term marital relationships” and “parasitic lifestyle” are simply inappropriate to youth; they typically have yet to experience an independent lifestyle characterized by marriage and self-support. Edens et al. (2001) also criticize items such as impulsiveness and irresponsibility as inappropriate markers for psychopathy in youth, again because adolescents are normatively more impulsive and stimulation seeking than adults. Furthermore, these researchers argue that throughout adolescence, youth continue to develop social perspective taking, which allows them to comprehend others’ points of views, time perspectives, and to consider the long-term consequences of their actions. In effect, youth cannot be expected to have the understanding of themselves required for mature behaviour and long term goal planning.

Edens et al. (2001) therefore identify the need for longitudinal research in order to determine both the applicability and capability of assessing PPD in youth.

Hart, Watt, and Vincent (2002) stress the continued major concerns regarding the linking of psychopathic traits in children and adolescents to the construct of psychopathy. They refer to the developmental literature in arguing that there are three central issues limiting the applicability of PPD to juveniles. First, juvenile psychopathy definitionally does not currently exist since, according to the Diagnostic and Statistical Manual, 4th edition (DSM-IV), an individual must be 18 years of age or older before a diagnosis of even the closely related antisocial personality disorder can be made. Essentially, there is a lack of consensus among developmental psychopathologists regarding the age at which personality becomes fully stable. In other words, if personality traits cannot be said to be stable during childhood or adolescence, then neither can personality *disorder* be stable during this developmental stage.

The second issue identified by Hart, Watt, and Vincent (2002) is that juvenile psychopathy may not resemble adult psychopathy, because of three critical developmental principles: heterotypic continuity, equifinality, and multifinality. These principles suggest that psychopathy is not likely manifested in similar patterns across the lifespan. Heterotypic continuity suggests that a trait may be expressed differently over the life course. Hart, Watt, and Vincent (2002) use the example of differential expressions of empathy; a child may display lack of empathy by having problems playing with peers, whereas an adolescent may display deficient empathy through problematic perspective taking. However, it is important to note that even though expressions of a trait may differ across the lifespan, the trait itself is still present, i.e., it is only its *expression* that is

unstable. The second and third principles are essentially the reverse of each other. Equifinality is the principle that a wide range of developmental pathways can result in the same outcome in adulthood, while multifinality represents the principle that the same original developmental pathway can result in a wide range of outcomes in adulthood. In the first instance then, adulthood psychopathy could be the result of a number of different distinct pathways that originate in childhood or adolescence. In contrast, multifinality suggests the possibility that many children and youth who display psychopathic traits in childhood or adolescence proceed along a different subsequent developmental path, resulting in alternative psychopathological outcomes, or even in normal development (Hart, Watt, & Vincent, 2002). Both of these possibilities highlight the importance of longitudinal research that will follow the developmental course of personality disorders, rather than research that relies upon the recall of life events and official documentation.

The final issue noted by Hart, Watt, and Vincent (2002) is that, even if psychopathy is a valid construct for children and youth, it may be impossible to assess reliably. Importantly, many measures of adolescent psychopathy, such as the PCL-YV, are simply downward extensions of adult measures that require a wide range of information provided by multiple sources. To score a PCL-R, for example, information should be taken not only from an interview conducted with the person of interest, but also from longstanding and extensive official records which document an individual's lifetime psychosocial adjustment. In contrast, when evaluating personality traits among samples of youth, while personal interviews can provide much needed information, there is typically a lack of sufficient documented information to corroborate interview information or to supplement the evaluation. In effect, the "very nature of the information

base used to assess juvenile psychopathy imposes a limit on the accuracy and reliability of the assessment” (Hart, Watt, & Vincent, 2002: 243). This issue, alone, emphasizes the importance of relying on research of psychopathy disorders in children and adolescence as opposed to using psychopathy as a diagnostic construct for clinical purposes. Without more research about personality disorders among children and youth in general, diagnoses of personality disorders such as psychopathy among youth should be made with extreme caution, if at all, given the potential for dangerous negative outcomes such as invalid labelling of youth.

Vincent (2002) identifies two key issues with respect to applying the construct of psychopathy to adolescents. Firstly, she identifies the absence of evidence supporting the long-term stability of psychopathic traits, and identifies the need for longitudinal research. Secondly, Vincent identifies a concern with respect to the measurement validity when assessing psychopathy in youth. In effect, Vincent proposes that it is very possible that characteristics of psychopathy are not manifested and, therefore, measured in the same manner in youth as with adults. In developing this argument, Vincent (2002) refers to the psychosocial literature that suggests that there may be an age-related measurement bias. As discussed, characteristics of impulsivity, irresponsibility, and grandiosity are common for many youth; as such, these characteristics may be inappropriate for use in assessing for psychopathic traits. Further, if this is the case, then it is likely that the prevalence of psychopathy among youth will be overestimated.

To test the theory that age-related measurement bias is present in assessments of psychopathy among youth, Vincent (2002) compared the PCL-R ratings for 444 adult male incarcerated offenders with the PCL:YV ratings for 269 adolescent male

incarcerated offenders. Using item response theory analyses, Vincent concluded that the 20-item model of psychopathy was inappropriate for use with youth. Specifically, Vincent cautioned that using the 20-item model will likely result in a higher rate of false positives, i.e. the number of youth being falsely identified as psychopathic. This overestimation is attributed to the presence of criminal behavioural traits. However, the results lent support to the use of the 13-item model in which reference to antisocial behaviours such as criminal offending were removed. Vincent concluded that the 13-item model of psychopathy was equally relevant in identifying psychopathic traits in both adults and youth. Further, Vincent concluded that the affective traits in particular appear to be present among youth at a relatively early age, and may be an identifiable precursor to the psychopathy construct as identified in adulthood. In sum, Vincent's analysis provided evidence for the presence of a coherent "psychopathy-related syndrome" in youth (2002: 80).

Lynam, Caspi, Moffitt, Loeber, and Stouthamer-Loeber (2007) are among the first set of researchers to provide longitudinal support for the construct of psychopathy as identified in youth. Over 250 youth who participated in the Pittsburgh Youth Study were assessed for psychopathic traits using the Child Psychopathy Scale (reported by mothers) at age 13 and were rated for psychopathy again at age 24 using the Psychopathy Checklist: Screening Version. The results of this study provided evidence for the moderate stability ($r = .31$) of psychopathy between adolescence and adulthood. This relationship was stronger for the Facet 3 (impulsive and irresponsible behavioural style) and Facet 4 (antisocial behaviours) dimensions of psychopathy ($r = .28$ and $r = .33$, respectively), but also held for the Facet 1 (interpersonal) and Facet 2 (affective) traits (r

= .19 and $r = .15$, respectively). The relationship between scores at age 13 and age 24 did not depend on the initial level of psychopathy; i.e. for psychopathy to be relatively stable between adolescence and adulthood, individuals did not have to score in the highest ranges of psychopathy initially. Further, psychopathy was able to provide incremental predictive validity over other explanatory variables for criminality recorded at age 13, i.e. race, family structure, family socioeconomic status, neighbourhood socioeconomic status, physical punishment, inconsistent discipline, lax supervisions, lack of positive parenting, peer delinquency, behavioural and cognitive impulsivity, verbal IQ, and previous delinquency.

There were, however, some limitations to this study. Of concern was the poor predictive power of psychopathy, in that most of those individuals at age 13 who were identified by the CPS as psychopathic did not receive a diagnosis of psychopathy again at age 24. Furthermore, psychopathy at age 13 only represented approximately 10% of the variance at age 24. These conclusions suggest that more longitudinal research is necessary to explore why psychopathic traits are stable in some but not others. Further, in referring to the stronger stability for Facets 3 and 4, Lynam et al. (2007) also note the necessity to explore whether and why personality traits differ in their particular stability over time.

Despite their cautions regarding the validity of the psychopathy construct in youth, Skeem and Cauffman (2003) also discuss research that has lent support to the predictive ability of the PCL:YV with respect to outcomes such as crime and violence. In addition, they refer to a fundamental assumption regarding the presence of a subgroup of young offenders who manifest the interpersonal, affective, and behavioural traits

identified among adult psychopaths. In other words, while many youth are impulsive, the impulsiveness of certain youth may be excessive or unusual, or may appear in combination with a range of other psychopathic features. Skeem and Cauffman (2003) further note that the subgroup of youth who are characterized by these traits will persist in their antisocial behaviours throughout adulthood. Most importantly, then, while several psychopathic traits may be common in adolescence, their degree of severity and the extent to which they occur in combinations supports an assessment of psychopathy using the PCL:YV. Similarly, Gretton, Hare, and Catchpole (2004) identify that it is relatively infrequent for youth to receive a high rating on the PCL:YV, given the extent of dysfunctional affective, interpersonal, and behavioural traits that are necessary to obtain a sufficiently high score. This thought is echoed by Hare (2003) who suggests that the characteristics of the psychopath, although possibly common to those found in adolescence, tend to be of a more serious or extreme nature than what is normally seen in adolescents.

Other researchers have cited developmental research in support of the application of the psychopathic construct to children and youth. Barry, Frick, and Killian (2003), in their research on the interpersonal dimension of psychopathy, state that children generally can accurately assess their self-worth, and that some children exaggerate and/or distort their self-view at an early age. Stouthamer-Loeber (1986) found that, while many young children experiment with deceit, pathological levels of deceit (i.e. chronological lying) are relatively uncommon; a small percent of children exhibit this trait.

Developmental research also indicates that many adult emotional states are present within very early stages of life, and that even from the toddler stage, children

develop their conscience by integrating a societal value framework. In effect, it should be possible to reliably assess dysfunctional development in early developmental stages (Salekin, 2006). Further, there is evidence that both normal and abnormal patterns of personality traits are relatively stable through late childhood into early adulthood (Gretton, Hare, & Catchpole, 2004). Crawford, Cohen, and Brook (2001) found evidence for the continuity of latent traits indicative of personality disorder (specifically, narcissistic, borderline, and histrionic) from early adolescence through early adulthood. Lynam, Caspi, Moffitt, Raine, Loeber, and Stouthamer-Loeber (2005) refer to the Five Factor Model of personality, which include the five big personality dimensions of Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness, arguing that traits of these dimensions, including those that are indicative of psychopathy, are identifiable in adolescence, and appear to be stable into adulthood.

In his book, *Without Conscience: The Disturbing World of the Psychopaths Among Us*, Hare (1993) argues that aspects of psychopathy are observable early, and that many parents of children later diagnosed as psychopathic were well aware that the related traits had existed during childhood. Hare (1993) refers to both clinical and anecdotal evidence that suggests that some psychopathic traits are present within a subgroup of children; pathological lying (lying constantly without thought or concern), defiance (of parents, teachers, rules), aggression (bullying, fighting), experimenting with sex at an early age, and hurting or killing animals are evident within early years of life. In addition, he notes that a callous indifference towards the thoughts and feelings of others is also evident at an early stage. Hare provides the example of “Jason”, a 13-year old youth who received one of the highest scores on the PCL among a subgroup of 13 to 18 year olds.

Since early in his life, Jason had been displaying traits that would be indicative of psychopathy in adulthood; engaging in serious crime (break and enter, theft, assault) since the age of six, and displaying a distinct lack of concern for the thoughts or feelings of his parents. Jason, the product of a stable home with professionals for parents, seemed to enjoy terrorizing them with his callous indifference and consistent engagement in a wide range of criminal activity (Hare, 1993).

Hare also provided anecdotal evidence of other children who, although not necessarily assessed as psychopathic, certainly displayed alarming behaviour and disturbing personality traits at an early age. He reported on the case of a nine year old who had his community at a loss as to how to respond as he raped and molested other children at knifepoint. Hare also recounted the case of “Tess”, a six and a half year old girl, who repeatedly terrorized her baby brother, punching him in the stomach at night, and bashing his head into the floor. Tess also admitted to stealing “big sharp [knives]” with the intent of killing her baby brother and adoptive mother (Hare, 1993: 171). These examples, albeit unusual cases of childhood and adolescence offending, support Hare’s assertion that “[p]sychopathy does not suddenly spring, unannounced, into existence in adulthood. The precursors ... reveal themselves early in life” (Hare, 1993: 157). Thus, it does appear that a construct similar to what is known as adult psychopathy is expressed in early childhood and adolescence. In effect, psychopathic traits are not suddenly expressed around the age of 18; on the contrary, the hypothesis is that there is typically a consistent long-term pattern of dysfunctional behaviours and traits evident from an early age. These long term patterns also are important in explaining the subtypes of psychopathy, as will be discussed in the next section.

Subtypes of Psychopathy

The construct of psychopathy itself is not necessarily a homogeneous concept. In fact, researchers have long discussed the presence of two fundamentally different forms of psychopathy: primary and secondary subtypes. Although said to have differing aetiologies, primary and secondary subtypes are essentially similar in their behaviours. This aetiological difference highlights the importance of determining whether there are varying subtypes of psychopathy. In recognizing that psychopathy is a heterogeneous concept, researchers and clinicians are recognizing that there are likely multiple causes of the disorder. This realization will ideally lead towards an exploration of the various possible developmental pathways that may lead to PPD.

Karpman (1941) initially distinguished between primary and secondary psychopaths; while phenotypically similar, these types differed in the origins and expressions of their affective symptoms. Karpman (1941) hypothesized that the affective deficit of primary psychopaths is constitutional, i.e. genetic, while the affective features of secondary psychopaths are caused primarily by childhood trauma. In other words, while all psychopaths are emotionally disturbed, such disturbances in secondary psychopaths are a reaction to major adverse early life experiences, such as parental rejection. Secondary psychopaths, therefore, possess the capacity for “higher human emotions” such as feeling empathy towards another; however, their adaptation to early childhood trauma (e.g. sexual abuse), such as closing off their feelings in order to protect themselves from emotional harm, results in the subsequent apparent inability to experience emotion (Karpman, 1941).

Lykken (1995) provided empirical and theoretical support for a biological distinction between primary and secondary psychopaths. First, he hypothesized that primary psychopaths are characterized by a low-fear quotient. According to Lykken (1995), there is an innate fear quotient present in all individuals, but it is minimal for primary psychopaths. He stated as well that “psychopaths” and “heroes” are both characterized by “low fear”. The major difference between a hero and a psychopath is the inadequate successful socialization of the latter, i.e. the absence of concern for punishment as disapproval (Lykken, 1995).

Lykken (1995) also relied on the biological model of personality (Gray, 1987; Gray & McNaughten, 1996) to differentiate between psychopaths in terms of the behavioural inhibition and activation systems. The Behavioral Activation System (BAS), which is associated with positive affect and impulsivity, activates the body’s response to items which it identifies as rewarding, whereas the Behavioral Inhibition System (BIS), which controls the body’s response to aversive stimuli and plays a role in negative affect, activates the response to threats of punishment and of non-reward. Imperfections in either of these systems can lead to different forms of psychopathic behaviour. For instance, primary psychopaths have been found to have a weak BIS, i.e., non-responsive to threats of punishment. This explains Lykken’s low-fear model where primary psychopaths have a “below-average endowment of innate fearfulness” (Lykken, 1995: 154). Further, in the absence of highly skilled parenting, primary psychopaths’ biological predisposition to low-fear arousal results in the failure of punishment to have a deterrent effect on antisocial behaviour. In contrast, secondary psychopaths are characterized by an

overactive BAS; while they may experience fear and anxiety similar to any normal individual, they are unable to control impulses to obtain a reward (Lykken, 1995).

Lykken's work integrating the BIS and BAS into explanations of psychopathy was recently supported by Newman, MacCoon, Vaughn, and Sadeh (2005), who compared scores of primary and secondary psychopaths against control participants using two measures of the BIS and BAS. The Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ) is a 48-item measure that assesses two subscales: Sensitivity to Punishment (SP), associated with BIS functioning, and Sensitivity to Reward (SR), associated with BAS functioning. In addition, the researchers used the BIS/BAS Scales, a 20-item questionnaire measuring sensitivity to aversive stimuli, sensitivity to rewards, motivation to achieve goals, and willingness to approach new stimuli. Newman et al. (2005) differentiated primary and secondary psychopaths using a combination of the PCL-R and a median-score split on the Welsh Anxiety Scale (WAS). In effect, primary psychopaths scored 30 or above on the PCL-R and 11 or less on the WAS. On the other hand, secondary psychopaths, who also scored 30 or above on the PCL-R, scored 12 or more on the WAS. Their findings with a sample of 517 incarcerated adult males indicated that primary psychopathy was associated with significantly lower BIS and SP scores than controls, but did not significantly differ with respect to their BAS or SR scores. In contrast, while secondary psychopathy was associated with significantly higher BAS and SR scores, the BIS was not significantly different between control participants and secondary psychopaths. An unexpected finding was that the secondary psychopaths also had significantly higher SP scores than controls, which the researchers explained as possibly due to increased levels of neuroticism (Newman et al., 2005).

Research has also documented that not all those who can be diagnosed as psychopathic engage in criminal behaviours. For instance, the interpersonal and affective traits of psychopathy have been defined as the true core features of psychopathy (e.g. Hare, Forth, & Strachan, 1992). Hall and Benning (2006) state that a subgroup of offenders who exhibit severe interpersonal and affective dysfunctions, but who do not also exhibit severe antisocial behavioural traits, i.e. “successful psychopaths”, can be identified. This type, while still acting in a manner that is distressing to others, either hide their criminal deviance, or have acted deviant in a non-criminal manner. This follows from Cleckley, who originally defined his psychopaths with reference to primarily personality traits, including only one indirect reference to criminality (“inadequately motivated social behaviour”) (Poythress & Skeem, 2006).

Using a categorical determination of psychopathy (i.e. psychopathic or not psychopathic) ignores the possibility that there are heterogeneous subtypes of psychopathic individuals. The presence of an arbitrary cut-off score that labels one as either psychopathic or not therefore clutters our understanding of the particular nature of the disorder. In other words, the effect of a dichotomous categorization of PCL scores can obscure the heterogeneity of psychopathic personality disorder.

The Psychopathy Checklist has been criticized for failing to differentiate between secondary and primary psychopaths. Although the PCL recognizes the presence of at least two, moderately correlated, factors of psychopathy, research suggests that it is unable to shed light on the heterogeneity of psychopathy. Research such as that discussed above suggests that psychopathy is a heterogeneous construct that may be associated with additional personality features unmeasured by the PCL, such as a lack of anxiety. Given

this, it is essential that future research attempt to further disaggregate the concept of psychopathy, recognizing additional clinically relevant traits that may serve to distinguish subtypes, in order to better understand the developmental sequence that leads to adulthood, and therefore lifetime, PPD.

Psychopathy: A Categorical or Dimensional Construct?

Despite studies such as Newman et al. (2005), subtype research is still in the early stages of development. However, Poythress and Skeem (2006) suggest that research with subtypes of psychopathy can inform the debate about whether psychopathy should be seen as a categorical disorder, as opposed to a dimensional one. While the early views of psychopathy, such as that proposed by Cleckley, spoke of a homogeneous group of individuals who could be categorized into a qualitatively distinct group, other views suggest that the disorder is better described by reference to a continuum.

Personality disorders in general have been recognized both as categorical and dimensional concepts. The DSM-IV conceptualizes personality disorders as categories; in assessing for the presence of personality disorder, evaluations are made on the basis of an absence/presence dichotomy, where the presence of a disorder is determined by surpassing a threshold or cut off score (e.g. Shedler & Westen, 2004; Ullrich & Marneros, 2004). Critiques of a categorical approach to personality disorder include that the cut off threshold at which a personality disorder is recognized as present is often arbitrarily determined, in addition to concerns regarding criteria overlap between different disorders (Ullrich & Marneros, 2004).

There is a great deal of debate regarding the appropriate underlying structure of the construct of psychopathy, and whether it is best referred to as a categorical or

dimensional construct. Harris, Rice, and Quinsey (1994) argue that PPD should be viewed as a taxon, or a discrete class. They refer to a taxon as something that is naturally occurring, whether or not it is acknowledged by others. Accepting that psychopathy is a taxon implies that this disorder is homogeneous, that all who are afflicted with it experience a similar degree of dysfunction.

Harris and colleagues classified 653 mentally disordered offenders using the PCL-R, although their ratings were limited as they were based solely on file reviews. They concluded that there is support for an underlying taxon of psychopathy (Harris, Rice, & Quinsey, 1994). However, Marcus, John, and Edens (2004) caution that while there was evidence of an underlying taxon for Factor 2, Factor 1 appeared to be continuous. In a subsequent re-analysis of their sample, Skilling, Harris, Rice, and Quinsey (2004) identified similarly that while the antisocial behavioural items appeared to have an underlying taxonic structure, the same could not be said for the personality-based characteristics of psychopathy. There are several possible explanations for these results; for instance, it could be that psychopathy is actually a dimensional construct. In contrast, it could be possible that the lack of interview data impeded the appreciation of affective and interpersonal personality traits and that psychopathy is actually a taxon.

Given the inconclusive nature of these previous results, Marcus, John, and Edens (2004) used the Psychopathic Personality Inventory (PPI) to analyze the personality traits of psychopathy. Their results suggested that the initial results were accurate, and that the personality features of psychopathy are best represented along a continuum. In other words, they found support for the dimensional model of psychopathy. Marcus and colleagues discuss several implications that a dimensional model of psychopathy would

have for clinicians and researchers. They note that often the aim of clinical psychopathy assessments is to determine either an absence or a presence of psychopathy. Similarly, researchers at times use a dichotomous measure of psychopathy to determine criminal justice outcomes such as recidivism following release into the community. However, given the nature of their results, Marcus et al. (2004) suggest that it may be more appropriate to interpret a psychopathy score as lying along a dimension of continuous scores.

This conclusion implies that as opposed to a qualitatively distinct category, psychopathic individuals can actually be described with reference to normal personality traits (Marcus, John, & Edens, 2004). It is presumed that humans have basic attributes that are necessary to normal personality functioning; however, for psychopathic individuals these attributes are impaired in some way (Marcus, John, & Edens, 2004). The causes of such impairment have yet to be concretely determined; however, as discussed previously, theory suggests that there may be genetic, environmental, or other biological causes. In addition, Marcus and colleagues note that a taxonomic approach to psychopathy would suggest a single cause or aetiology; however, with reference to a dimensional model, multiple aetiologies become possible (Marcus, John, & Edens, 2004).

As stated, in contrast to a categorical or taxonomic approach, dimensional models of personality allow for the recognition of a continuum of criteria (Shedler & Westen, 2004). Hare's original operationalization of psychopathy in the PCL and PCL-R categorized the traits of psychopathy into two main factors: the personality features, located in the Interpersonal/Affective Factor, and the antisocial behavioural features, located in the Social Deviance Factor. Theoretically, psychopathy can be described by

reference to these two main dimensions, each of which refers to a subset of traits. As such, the PCL-R model of psychopathy is a dimensional model of personality dysfunction, as it provides a continuum of scores along which an individual can fall, allowing for a more finite degree of comparison with the prototypical psychopath (Ogloff, 2006). However, the PCL-R can also produce a categorical determination; by using a defined cut point of a score of 30 out of 40, researchers and clinicians can speak to the absence or presence of psychopathy.

Vasey and colleagues (2005) argue that the debate regarding categorical versus dimensional models of psychopathy can be clarified through increased knowledge regarding the underlying structure of the disorder, which will reveal whether there is a distinct group of individuals who can be identified as “psychopathic”. A recently developed model of psychopathy, called the Comprehensive Assessment of Psychopathic Personality Disorder (CAPP), attempts to clarify this underlying structure with reference to a comprehensive set of symptoms said to be indicative of PPD. The CAPP brings the discussion of psychopathy fully back into the realm of personality. While the PCL-R assesses for personality features of psychopathy, it also includes reference to antisocial behaviours such as criminal offending or substance misuse. However, some researchers have taken the position that antisocial behaviours are best treated as consequences of these personality traits, rather than as symptoms of PPD (e.g. Cooke, Michie, Hart, and Clark, 2004). As such, the CAPP, developed by David Cooke, Stephen Hart, and Caroline Logan, proposes a six-dimensional model of psychopathy that is fully reflective of personality traits.

The Comprehensive Assessment of Psychopathic Personality Disorder

The Comprehensive Assessment of Psychopathic Personality Disorder (CAPP) represents a model of psychopathy that is characterized by the hierarchical construct of psychopathy as defined by reference to six domains of functioning. It is proposed that the CAPP represents the full syndrome of psychopathy. Furthermore, the CAPP measures personality traits with respect to both lifetime, as well as recent trait expressions. In other words, the CAPP has been developed to allow for a determination of stability versus change. The CAPP can be used in clinical settings, for instance, to determine whether a recent treatment experience has affected dysfunctional traits. The CAPP therefore also lends itself to longitudinal research, in which the stability of traits over time can be determined.

The PCL has been criticized by others for its failure to include reference to clinically relevant traits such as a lack of anxiety, as well as for its over-reliance on criminal behaviours as an indicator of PPD. For instance, Cooke and Michie (2001) have asserted the need to add traits drawn from other essential clinical indicators of psychopathy, such as emotional coldness, incapacity for love, fearlessness, and anxiety. Lilienfeld and Andrews (1996) also argue that there is a need to further clarify the boundaries of the psychopathic construct. They have developed the Psychopathic Personality Inventory (PPI), designed to assess the large number of personality dimensions relevant to psychopathy. The PPI is over-inclusive with respect to potentially relevant traits; traits indicative of fearlessness, the inability to form close attachments, low ambition, authority problems, and a lack of anxiety all were identified as clinically relevant to psychopathy (Lilienfeld & Andrews, 1996).

In developing the CAPP, Cooke, Hart, and Logan (unpublished manuscript) reviewed the available literature and research on PPD and similar constructs, such as sociopathy and antisocial personality disorder. In addition, they conducted interviews with clinicians from a multitude of theoretical backgrounds who had experience with PPD-patients. Based on their experiences, clinicians were asked to identify which symptoms they identified as central to psychopathy. Through this process, Cooke, Hart, and Logan determined that the construct of PPD is seriously underconceptualized in terms of the inclusion of relevant symptoms. Cooke, Hart, and Logan (unpublished manuscript) subsequently created the Comprehensive Assessment of Psychopathic Personality Disorder (CAPP), extending the current conceptualization of PPD to a six-dimensional hierarchical model of personality.

This model re-conceptualizes psychopathy by broadening the boundaries of our understanding of this disorder and presenting it as a hierarchical model in which PPD is a superordinate factor, represented by heterogeneous symptom clusters. By including domains representative of dominance, attachment, and self-concept, their model is consistent in fitting with other major theories of personality, and, therefore, integrates theories of psychopathy with historical theories of abnormal and normal personalities (Cooke, Hart, & Logan, unpublished manuscript). A brief discussion of the latter theories is important in understanding their model.

Personality shapes our behaviour; it defines who we are and it influences what we do and how we interact with others. In effect, how we interact with ourselves, i.e. our self-concept and how it is impacted by introjection and others reflects our personality traits and interpersonal style (Greenberg & Mitchell, 1983). Accordingly, dysfunctional

or socially devious behaviours are a direct reflection or expression of personality. There are several models of personality that attempt to describe this complex phenomenon. Interpersonal models usually reference an interpersonal circle, or circumplex, shaped by two key dimensions of power/control and affiliation (e.g. Dolan & Blackburn, 2006; Blackburn, Logan, Renwick, & Donnelly, 2005). These dimensions are orthogonal; they are mutually independent and characterized by opposite poles. The opposing poles of dominance and submission anchor the dimension of power/control, while the opposing poles of hostility and nurturance/friendliness anchor the dimension of affiliation. According to the interpersonal circle model, personalities are distinguishable with reference to these two dimensions. Thus, a person can be described as friendly-submissive, hostile-dominant, friendly-dominant, or hostile-submissive. Blackburn et al. (2005) argue that the hostile-dominant interpersonal lifestyle explains psychopathy. This model of personality is useful, yet simplistic, since it refers to only two main dimensions which shape all of personality. Yet not all people who are dominant and friendly should be located within the same quadrant, as they likely differ in other important ways, for example, with respect to their self-concept or their cognitive processes. Arguably then, there are additional personality dimensions which are relevant in shaping as a person, how one behaves towards others.

The model of personality structure presented by Cooke, Hart, and Logan (unpublished manuscript) is a dimensional model of personality which consists of the combination of six dimensions of personality dysfunction characteristic of PPD. The six-dimensional personality structure is hierarchically represented by the personality disorder of psychopathy. In other words, this complex personality disorder is a multi-dimensional

series of varying combinations of dysfunctional personality traits. In comparison, Hare proposes that psychopathy consists of only affective, interpersonal, and behavioural dysfunctional traits. Cooke et al. (unpublished manuscript) after reviewing the literature on PPD and personality theories along with interviewing experienced clinicians, conclude that PPD as currently conceptualized is lacking the key personality components of self-concept, attachment styles, and cognitive processes.

As discussed above, while recognizing the presence of three “traditional” dimensions of interpersonal, affective, and behavioural functioning, Cooke et al. identified an additional three dimensions of clinical relevance to PPD; attachment, emotional, and self-functioning. Accordingly, the prototypical psychopath is characterized as having dysfunctional interpersonal relationships involving attempts to dominate; displaying shallow and inconsistent emotions; acting impulsively and irresponsibly; being unable to fully commit to another person; exhibiting a high degree of suspiciousness and dysfunctional attribution processes; and maintaining a distorted view of the self and one’s own capabilities. In contrast, according to the PCLs conceptualization of psychopathy based on the most recent four-factor model, the prototypical psychopath is characterized by a more limited description: domination of interpersonal relationships, displaying shallow and inconsistent emotions, and acting impulsively and irresponsibly, often engaging in a wide range of criminal behaviours.

While the CAPP refers to behavioural actions indicative of the presence of a dysfunctional symptom, the CAPP does not include non-clinically relevant items such as criminal history. Again, it is important to reiterate, first, that one of the fundamental criticisms of the PCL was that it was tautological when it was used to predict past

criminal behaviour since indicators of past criminal behaviours were utilized as part of the behavioural dimension of psychopathy. Secondly, this behavioural dimension too often was the sole significant predictor of future criminal behaviour, when this dimension is not the central defining one for psychopathy. The affective and interpersonal dimensions are, and remain, central to any conceptualization of psychopathy.

Hare's PCL instruments have contributed considerably to our initial understanding of PPD, as well to its theoretical and empirical associations with general and violent offending. However, the limitations of the PCL justify PPD being re-conceptualized into a fully represented model of personality dysfunction. With their new measure of PPD, Cooke et al. (unpublished manuscript) have broadened the conceptualization of the psychopathic individual and, arguably, provide a more valid model of psychopathy.

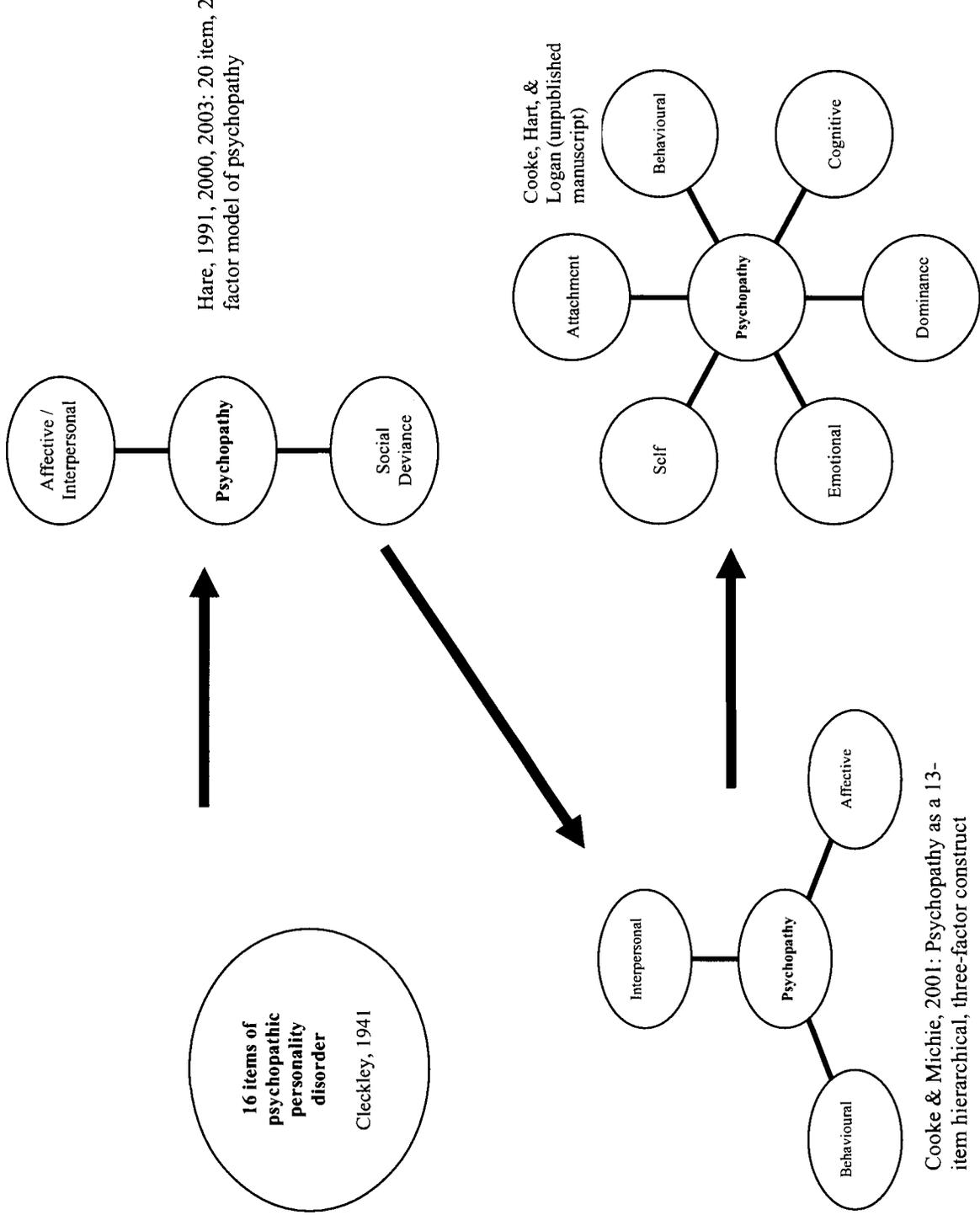
The CAPP Model of Psychopathy

Over the past several decades, our conceptualization of PPD has broadened, beginning with a description of the disorder and its common traits and leading today into a hierarchical multiple-dimension model of personality disorder (see Figure 1). Our understanding of PPD was greatly influenced by the work of Hervey Cleckley who, in the mid-1900s, provided rich clinical descriptions of the disorder based on his experience with PPD-afflicted clients. Cleckley described the prototypical psychopath with reference to 16 key items (see Table 1).

Table 1: Cleckley's 16 Item Model of Psychopathy (adapted from Cleckley, 1964, pp: 362-363).

Item	Item Description
1.	Superficial charm and good "intelligence"
2.	Absence of delusions and other signs of irrational thinking
3.	Absence of "nervousness" or psychoneurotic manifestations
4.	Unreliability
5.	Untruthfulness and insincerity
6.	Lack of remorse or shame
7.	Inadequately motivated antisocial behaviour
8.	Poor judgment and failure to learn by experience
9.	Pathologic egocentricity and incapacity for love
10.	General poverty in major affective reactions
11.	Specific loss of insight
12.	Unresponsiveness in general interpersonal relations
13.	Fantastic and uninviting behaviour with drink and sometimes without
14.	Suicide rarely carried out
15.	Sex life impersonal, trivial, and poorly integrated
16.	Failure to follow any life plan

Figure 1: Development in the Models of Psychopathy



Psychopathy was first operationalized by Dr. Robert Hare, who constructed the original Psychopathy Checklist through reference to Cleckley's writings. Hare's original conceptualization and operationalization of the disorder consisted of a global 7-point rating scale along which clinicians would determine the extent to which the individual being assessed fit the model of the "prototypical psychopath" elucidated by Cleckley (Hare & Neumann, 2006). Given that clinicians made their own determinations of what evidence would be used to apply a particular rating, Hare recognized the need for a more standardized assessment tool that could be used more objectively by raters. Hare subsequently reduced the criteria included on the global assessment to a 22-item instrument that measured the construct of psychopathy. With the removal of the items "drug or alcohol abuse not direct cause of antisocial behavior" and "previous diagnosis as psychopath or similar", the instrument was further reduced into 20-items (Hare & Neumann, 2006).

As discussed previously, Hare's 20 items were factored into a two-factor model of psychopathy (see Table 2) characterized by reference to affective/interpersonal traits, as well as antisocial behaviours. The extent to which the antisocial behaviour items of psychopathy play a relevant role in indicating the presence of the disorder has been debated by many researchers. Most importantly, Cooke and Michie (2001) re-conceptualized the Hare's description of PPD by removing seven items that referred to antisocial behaviours that Cooke and Michie regarded as inappropriate for inclusion. Cooke and Michie's (2001) reconceptualization of psychopathy led to a three-factor model represented by 13 items (see Table 2).

Table 2: Hare's 20-item and 13-item Models of Psychopathy (adapted from Hare & Neumann, 2006 and Corrado, Vincent, Hart, & Cohen, 2004)

Item	Item Description	Factor Number (2 Factor Model)	Factor Number (3 Factor Model)
1	Glibness / Superficial charm	1	1
2	Grandiose sense of self worth	1	1
3	Need for stimulation / Prone to boredom	2	3
4	Pathological lying	1	1
5	Conning / Manipulative	1	1
6	Lack of remorse or guilt	1	2
7	Shallow affect	1	2
8	Callous / Lack of empathy	1	2
9	Parasitic lifestyle	2	3
10	Poor behavioral controls	2	†
11	Promiscuous sexual behavior	†	†
12	Early behavioural problems	2	†
13	Lack of long-term goals	2	3
14	Impulsivity	2	3
15	Irresponsibility	2	3
16	Failure to accept responsibility	1	2
17	Short-term marital relationships	2	†
18	Juvenile delinquency	2	†
19	Revocation of conditional release	†	†
20	Criminal versatility	†	†

* Factor 1 = Interpersonal/Affective, Factor 2 = Behavioral

** Factor 1 = Interpersonal, Factor 2 = Affective, Factor 3 = Behavioural

† Not included in model

The current model of psychopathy, as described by Cooke, Hart, and Logan (unpublished manuscript) develops the conceptualization of psychopathy to an even broader representation. By referring to six domains of functioning, each of which is characterized by multiple symptoms or items, Cooke and colleagues provide a model of psychopathy that allows for the determination of heterogeneous clusters (Table 3).

Table 3: Cooke, Hart, and Logan (unpublished manuscript) 6-Dimensional Model of Psychopathy (the CAPP Domains, Symptoms, and Illustrative Indicators)

<i>Domain</i>	<i>Symptom</i>	<i>Illustrative Indicator</i>
Attachment	A1 <i>Detached</i>	Remote, Cold, Distant
	A2 <i>Uncommitted</i>	Unfaithful, Undevoted, Disloyal
	A3 <i>Unempathic</i>	Uncompassionate, Cruel, Callous
	A4 <i>Uncaring</i>	Inconsiderate, Thoughtless, Neglectful
Behavioural	B1 <i>Lacks Perseverance</i>	Idle, Undisciplined, Unconscientious
	B2 <i>Unreliable</i>	Undependable, Untrustworthy, Irresponsible
	B3 <i>Reckless</i>	Rash, Impetuous, Risk-Taking
	B4 <i>Restless</i>	Overactive, Fidgety, Energetic
	B5 <i>Disruptive</i>	Disobedient, Unruly, Unmanageable
	B6 <i>Aggressive</i>	Threatening, Violent, Bullying
Cognitive	C1 <i>Suspicious</i>	Distrustful, Guarded, Hypervigilant
	C2 <i>Lacks Concentration</i>	Distractable, Inattentive, Unfocused
	C3 <i>Intolerant</i>	Narrow-minded, Bigoted, Hypercritical
	C4 <i>Inflexible</i>	Stubborn, Rigid, Uncompromising
	C5 <i>Lack Planfulness</i>	Aimless, Unsystematic, Disorganized
Dominance	D1 <i>Antagonistic</i>	Hostile, Disagreeable, Contemptuous
	D2 <i>Domineering</i>	Arrogant, Overbearing, Controlling
	D3 <i>Deceitful</i>	Dishonest, Deceptive, Duplicitous
	D4 <i>Manipulative</i>	Devious, Exploitative, Calculating
	D5 <i>Insincere</i>	Superficial, Slick, Evasive
	D6 <i>Garrulous</i>	Glib, Verbose, Pretentious
Emotional	E1 <i>Lacks Anxiety</i>	Unconcerned, Unworried, Fearless
	E2 <i>Lacks Pleasure</i>	Pessimistic, Gloomy, Unenthusiastic
	E3 <i>Lacks Emotional Depth</i>	Unemotional, Indifferent, Inexpressive
	E4 <i>Lacks Emotional Stability</i>	Temperamental, Moody, Irritable
	E5 <i>Lacks Remorse</i>	Unrepentant, Unapologetic, Unashamed
Self	S1 <i>Self-Centred</i>	Egocentric, Selfish, Self-Absorbed
	S2 <i>Self-Aggrandizing</i>	Self-Important, Conceited, Condescending
	S3 <i>Sense of Uniqueness</i>	Extraordinary, Exceptional, Special
	S4 <i>Sense of Entitlement</i>	Demanding, Persistent, Sense of being Deserving
	S5 <i>Sense of Invulnerability</i>	Invincible, Indestructible, Unbeatable
	S6 <i>Self-Justifying</i>	Minimizing, Denying, Blaming
	S7 <i>Unstable Self-Concept</i>	Labile, Incomplete, Chaotic Sense of Self

The CAPP delineates a six-dimensional hierarchical model of PPD. Each of the six domains is characterized by impaired functioning in a particular area (see Table 3). The *Attachment* domain is defined in terms of both the social milieu and attachment to others. Attachment conveys an emotional bond with another person. Dysfunctional attachment-styles, therefore, are measured by the degree of affiliation, intimacy, and acceptance by others that an individual will attempt to achieve in their interpersonal interactions. The domain is characterized by *detached* (A1), *uncommitted* (A2), *unempathic* (A3), and *uncaring* (A4) symptoms. These symptoms can subsequently be described by adjectives including coldness, disloyalty, callousness, and thoughtlessness. The domain of attachment was not directly assessed by prior measures of psychopathy such as the PCL. Attachment instead would likely have been represented by the combination of interpersonal and affective traits, as these two dimensions of personality would likely regulate the extent to which one is able to participate in normal committed relationships in which one experiences an emotional connection with another.

The *Behavioural* domain is characterized by problems in organizing goal-directed activities. It includes traits such as impulsivity and sensation seeking, and its symptomatology is comprised of *lacking perseverance* (B1), *unreliability* (B2), *recklessness* (B3), *restlessness* (B4), *disruptiveness* (B5), and *aggressiveness* (B6). Key adjectival descriptors include idleness, irresponsibility, impetuousness, fidgetiness, unmanageability, and bullying. The behavioural domain is similar to the behavioural domain as measured by the PCL, as it includes reference to traits such as impulsivity and irresponsibility. However, the behavioural domain on the CAPP does not include symptoms indicative of antisocial behaviours such as criminality. Instead, it endeavours

to explore the personality traits and symptom constructs that may underlie and promote such behaviours.

The *Cognitive* domain focuses on mental flexibility and adaptability. This domain includes problems such as distractibility, intolerance, and suspiciousness. In addition, it is concerned with the process by which a person will focus, maintain, and allocate attention, encode and process information, organize their thoughts, as well as their attributional process. The domain includes the following five symptoms – *suspicious* (C1), *distractible* (C2), *intolerant* (C3), *inflexible* (C4), and *lacks planfulness* (C5), which are characterized by adjectival descriptors including hypervigilance, inattentiveness, narrow-mindedness, stubbornness, and disorganization. Cognitive abilities were not directly assessed by the PCL, although the inability to form future plans was an original item included on the PCL. The CAPP goes further in providing information on cognitive strategies than does the PCL as it refers to the maladaptive ways in which information is processed and interpreted. It allows for the exploration of hostile attributions and suspicious thoughts, traits that were not accessible directly by the PCL.

The *Dominance* domain is characterized by interpersonal difficulties, such as undue assertiveness and excessive status-seeking. This domain is concerned with the extent of power or control the individual attempts to take in interpersonal relations. The specific symptomatology of the dominance domain includes *antagonistic* (D1), *domineering* (D2), *deceitful* (D3), *manipulative* (D4), *insincere* (D5), or *garrulous* (D6) traits. These six symptoms can be assessed along a dominance-submissiveness contingency, which indicates to what extent adjectival descriptors such as contemptuousness, arrogance, deceptiveness, deviousness, superficiality, or

pretentiousness appear evident. The dominance domain is essentially an extension of the interpersonal domain as measured by the PCL. The CAPP in this sense is similar to the PCL, in that it refers to traits such as being manipulative or garrulous. However, the CAPP goes further in providing a broader exploration of attempt to dominate and instigate others.

The *Emotional* domain is conceptualized as the physiological response to persons and situations; thus, impairments will involve the experience and manifestation of shallow, labile emotions and mood dysregulation. This domain is assessed by the consideration of tone, depth, and appropriateness of affective responses. Its symptomatology involves the lack of: *anxiety* (E1), *pleasure* (E2), *emotional depth* (E3), *emotional stability* (E4), and *remorse* (E5). These symptoms are described by the following adjectives: fearlessness, pessimism, indifference, irritability, and unapologetic (Cooke, Hart, & Logan, unpublished manuscript). The emotional domain is related to the affective domain of the PCL in that both refer to emotional capacity. Both domains include similar traits such as lack of emotional depth and lack of remorse.

Lastly, the *Self* domain encompasses problems with the individual's identity or individuality, and thus its symptomatology includes self-centered or self-aggrandizing behaviours. This domain measures the extent to which the individual has an accurate awareness of their own identities, which includes their recognition of their personality traits, abilities, qualities, and desires. Problems in this domain result from the incongruence between one's own knowledge and awareness about the self as compared to other's expectations and preferences. It also results from the inability of the individual to set potential goals and imagine possible futures, due to their lack of awareness regarding

their abilities and interests. The domain thus includes seven symptoms – *self-centered* (S1), *self-aggrandizing* (S2), *sense of uniqueness* (S3), *sense of entitlement* (S4), *sense of invulnerability* (S5), *self-justifying* (S6), and *unstable self-concept* (S7), which can be described by adjectives such as egocentricity and selfishness, conceit, sense of exceptionality, demandingness, sense of invincibility, denial, or incompleteness, (Cooke, Hart, & Logan, unpublished manuscript). Including reference to the self concept in the CAPP allows one to determine the extent to which one's identity plays a role in their interactions with others, something which was not assessed using PCL measures of psychopathy.

In addition to the inclusion of these new dimensions relevant to PPD, the CAPP also broadens the conceptualization of personality dimensions by including multiple symptoms characteristic of each domain of functioning, each of which are rated independently. For example, the PCL measures the interpersonal traits of psychopathy using the absence, possible presence, and definite presence of the Impression Management, Grandiose Sense of Self Worth, Pathological Lying, and Conning/Manipulative items. In contrast, the Interpersonal domain of the CAPP (labelled Dominance) measures interpersonal dysfunction in relation to the six symptoms of *Antagonism*, *Domineering*, *Deceitful*, *Manipulative*, *Insincere*, and *Garrulous*, each of which are additionally represented by a series of illustrative and behavioural indicators.

The CAPP represents a hierarchical dimensional model of personality dysfunction. In contrast, categorical models result in the simple distinction between absence and presence of a dysfunctional trait. In other words, dimensional models allow for the determination of not only whether a trait appears to be dysfunctional, but also the

extent of dysfunction. However, as this measure has never before been applied in research, the current study will take the first steps towards validating its use in assessing PPD, by establishing the interrater reliability of the instrument.

VALIDITY AND RELIABILITY

Research conducted in the fields of social science (e.g. criminology, psychology) is based upon the use of observations and the use of measurements. Psychometrics refers to the measurement of psychological variables such as personality traits. There is an obvious need to ensure that all data collected in the course of research is as accurate as possible; as such, reliability and validity assessments are a cornerstone of social science research. Simply put, reliability refers to the extent to which an instrument produces the same measure each time it is applied (Everitt, 1996). Validity involves an assessment of the extent to which an instrument truly is assessing the construct of interest (Everitt, 1996). Both are necessary in conducting research. Reliability essentially involves the stability of a measurement instrument or method; but while an instrument may be reliable, it is not necessarily also accurate (Maxfield & Babbie, 2005). As such, it is also necessary to assess the extent to which a measurement instrument or method is valid, as this will support the accuracy of the measure used (Maxfield & Babbie, 2005). There are many forms of both reliability and validity; these will be discussed in the following section.

Validity

Validity assessments are an integral aspect of the research process. Whereas external validity refers to the extent to which results obtained from a particular sample can be applied to the larger population of interest, internal validity refers to the extent to which a measure can be said to be an accurate measure of the phenomenon it is designed to assess (Vogt, 2007). Therefore, internal validity essentially means that the concept of

interest is being accurately measured by the research instrument (Maxfield & Babbie, 2005). There are several forms of internal validity that must be assessed in order to determine that the results obtained are accurate.

In order to determine whether on the face of things, an instrument is assessing the construct of interest, face validity must be assessed. Face validity simply refers to the extent to which the instrument is a sensible measure of the construct of interest. An instrument that asks participants to indicate the frequency with which they drive while impaired by alcohol (e.g. never, sometimes, often) would likely be seen as having face value in measuring impaired driving behaviours while simply counting the number of cars leaving a bar parking lot at closure would not. In other words, the former measure would likely be viewed as an indicator of the concept of interest (impaired driving behaviours) while the latter would not (Jackson & Verberg, 2007).

The validity of an instrument can also be tested by determining the extent to which an instrument can accurately predict scores on another similar, already validated instrument (Maxfield & Babbie, 2005). This is known as criterion-related validity. Criterion-related validity can be further divided into predictive and concurrent validity. In both cases, the question is the extent to which a measurement instrument is closely related to something that it ought to relate to, i.e. the criterion of interest (Vogt, 2007). Concurrent validity compares whether two measures that occur at relatively the same time can be said to measure the same criterion, whereas predictive validity refers to the ability of a measurement instrument to correlate with a future outcome (Vogt, 2007). For instance, Lee, Vincent, Hart, and Corrado (2003) explored the concurrent validity of the Antisocial Process Screening Device, a self-report measure of psychopathy (Frick &

Hare, 2001), through comparing it with psychopathy assessments made using the PCL:YV. Given that the APSD is essentially a downward extension of the PCL criteria, the researchers expected it to measure psychopathy similarly when compared to the PCL:YV assessments. However, Lee and colleagues (2003) failed to establish the concurrent validity of the APSD as their results suggested that the APSD does not appear to assess the personality characteristics of psychopathy.

Predictive validity has often taken the form of predicting future criminal recidivism using PCL:YV total scores. Lee et al. (2003) also assessed the predictive validity of the APSD by testing its ability to predict psychopathy diagnoses as determined by the PCL:YV. The researchers failed to establish the predictive validity of the APSD as the instrument was unable to predict high scores on the PCL:YV.

An abstract concept can be defined in many ways, and establishing the validity of the content is another manner in which validity of an instrument can be supported. Content validity essentially refers to the extent to which the instrument assesses the full definition of the phenomena of interest. Content validity belies the importance of a theoretical framework that underlies the measurement tool itself. By theorizing about all possible content that may be related to the phenomenon of interest, an instrument can be built to reflect upon all such content. In order to determine whether an instrument has achieved content validity, Vogt (2007) recommends the use of panels, by which expert judges can assess the relevance of items that are said to represent the construct and determine whether the items presented represent the fully array of content underlying the construct of interest.

As an example, Burns (2000) refers to the Psychopathy Screening Device (now known as the Antisocial Process Screening Device) in arguing that items reflected in an instrument must be specific to the construct of interest; in other words, that the content of the construct is clearly defined and specific. According to Burns, the Psychopathy Screening Device fails to establish content validity as there is a great deal of overlap between its content and that found in measures of Attention Deficit Hyperactivity Disorder, Oppositional Defiant Disorder, and Conduct Disorder; therefore, the construct of psychopathy, according to Burns (2000), has not been clearly elucidated by the Psychopathy Screening Device.

Construct validity helps to establish the validity of an instrument by measuring phenomena that are theoretically associated with the phenomena of interest. There are two forms of construct validity; these are convergent and discriminant validity. Convergent validity is measured by determining the extent to which a measurement instrument positively correlates with another that is said to measure the same construct of interest (Vogt, 2007). In contrast, discriminant validity is supported by showing that an instrument is negatively correlated with one that is said to measure a different construct of interest. Discriminant validity assesses the extent to which a high score on the measurement instrument is related to a low score on a test that is said to measure the theoretical opposite of the construct of interest (Vogt, 2007).

Reliability

The validity of an instrument is somewhat dependent upon the reliability of an instrument. An instrument cannot be said to be an accurate measure of a construct if the measure fails to consistently produce similar results when expected to (Vogt, 2007).

Reliability refers to the replicability of an instrument: the extent to which an instrument produces similar results when applied to the same construct over time, or the extent to which different individuals who study the same phenomenon using the same measurement tool come to the same conclusion (Vogt, 2007). Unreliability in measurement results in errors in the assessment of variables, and decreases the chances that one will find a significant relationship between variables of interest when such a relationship truly exists (Vogt, 2007).

Reliability problems can arise as the result of the instrument itself, or from those administering it. For instance, the manner in which a test or instrument, such as a breathalyzer, is administered by various individuals, such as several police officers, can differ, producing varying results (Maxfield & Babbie, 2005). Alternatively, whenever one person is a source of information, there is the chance that their responses are less reliable due to the subjectivity of their judgment. For instance, when one individual is interpreting information, Maxfield and Babbie (2005) caution that it is difficult to determine accurately how much of that information is reflective of true variation in the phenomenon of interest as opposed to due to the unique interpretation of the individual.

The test-retest method is commonly used to assess the former problem. A test is reliable to the extent that it provides similar information when it is repeatedly applied to the same phenomenon. If an individual is administered a test at an initial test period, and then is retested six months later, and if the results are not expected to vary (for instance, as the result of some intervention that has been given during this period), then the extent to which the results *do* vary is indicative of an unreliable test (Maxfield & Babbie, 2005).

The reliability of the items on the test may also be of issue. Internal consistency reliability refers to the extent to which items on a scale that are said to measure different aspects of the same concept produce reliable measures of that concept (Vogt, 2007). For instance, a scale measuring intelligence may be composed of six different items that all measure a different aspect of intelligence. In determining whether it is appropriate to treat these six items as a single index of intelligence, reliability coefficients must be used. Cronbach's alpha is a statistical analysis that is often used to assess the degree of internal consistency reliability of these items. Cronbach's alpha is essentially a correlation coefficient between different items on a scale; if the reliability coefficient produced is .00, it indicates a complete lack of consistency between the items and suggests that they are not reliably measuring the same concept. If the reliability coefficient produced is 1.00, then the items are said to be correlated together perfectly. Typically, acceptable levels of Cronbach's alpha range around .70 and higher (Vogt, 2007).

When measurement instruments require that raters make subjective assessments, the possibility of measurement error is introduced. Measurement is essentially the process by which the abstract world of concepts and constructs is related to empirical or measurable indicators. When engaging in measurement, it is possible that the true value of the phenomenon is not recorded, for instance, due to systematic or random error. Systematic error refers to the tendency to systematically over- or under-rate a concept (Jackson & Verlag, 2007). For instance, a person may consistently report their height as two inches taller than it really is; this tendency is said to be systematic because it is not related to random fluctuations in estimates. Random error, then, refers to the fluctuation that *does* occur around the actual true value (Jackson & Verlag, 2007). With respect to

random error, it is equally possible that a person will overestimate or underestimate their response; they do not do so consistently, but instead by chance. For example, a judge may at times overestimate the chance that a sexual offender will re-offend, while at other times they may underestimate this chance. The judge does not necessarily have a tendency to always overestimate or underestimate the chance of sexual offenders re-offending; instead, their estimations tend to fluctuate randomly.

Thus, when individuals themselves are making judgments or evaluations, there is a concern regarding the extent to which the raters themselves are reliable. This form of reliability is known as interrater reliability, and it involves the extent to which, for example, two individuals given the same information will interpret that information in similar ways (Maxfield & Babbie, 2005). The extent to which raters are unreliable in assessing the same phenomenon using the same measure suggests either that the measure itself is unreliable in producing consistent results, or that the raters themselves are unreliable in applying the measure the same way.

While measurement error is common in the social sciences, instruments that involve human judgment are particularly at risk, due to the subjectivity of human interpretation (Shrout & Fleiss, 1979). It is essential to estimate the extent of such measurement error to determine the extent to which interviewer judgments and interpretations can be seen as reliable. To do so, an interrater reliability index must be calculated (Shrout & Fleiss, 1979). One method in which consistency between raters can be detected is by providing raters with a sub-sample of the phenomena of interest and having them independently code the same sample (Maxfield & Babbie, 2005). In effect, conducting interrater reliability analysis with continuous measures involves correlating

the scores independently assigned by two raters to the same phenomenon. One such measure of interrater reliability is the intraclass correlation coefficient (ICC).

Intraclass Correlation Coefficients

An ICC represents the amount of observed variance that is due to true variability between subjects (Everitt, 1996). ICC values may range between -1.00 and +1.00; stronger interrater reliability will be evident as scores approach +1.00, indicating that there is little variation between raters. When this is the case, the variation that *is* observed may be more confidently attributed to true differences in the phenomena of interest.

Interrater reliability assessments have been reported for both the PCL-R and the PCL:YV. However, it is somewhat difficult to compare interrater reliability directly across studies as, while the methods used to determine reliability often rely upon ICC analyses, the model employed, for instance, whether a two-way mixed effects model or a two-way random effects model, is infrequently reported (e.g. Edens et al., 2006). Further, the nature of the ICC that is reported varies. While most studies have reported only the total score reliability (e.g. Dolan & Rennie, 2006; Jones, Cauffman, Miller, & Mulvey, 2006), ICCs for both instruments have also been determined at both the total and factor levels. Other studies have reported factor scores as opposed to the total score. For instance, in a study of 157 youth referred to Youth Forensic Psychiatric Services in British Columbia, Gretton, Hare, and Catchpole (2004) reported ICCs for Factor 1 (.72) and Factor 2 (.68), but failed to report the ICC for the total score.

Reliability as reported for the PCL-R and PCL:YV has typically been quite high. As is shown in Table 4, sample interrater reliability, as assessed using ICCs for the PCL measures total scores has ranged from .79 (PCL) to .98 (PCL:YV). Skeem and Cauffman

(2003) similarly report that total score interrater reliability of the PCL:YV has ranged from ICC =.80 to ICC=.93.

Table 4: Sample ICCs for Total Score Interrater Reliability of PCL Measures

Author(s)	Sample	Instrument	ICC
Gretton et al. (2004)	Files of Canadian young offenders	PCL:YV	N/A
Vincent (2002)	U.S. county inmates	PCL	.79
Vincent (2002)	U.S. federal inmates	PCL	.84
Lynam et al. (2007)	Youth in the Pittsburgh Youth Study	PCL:SV	.86
Salekin et al. (2006)	American child and adolescent offenders	PCL:YV	.89
Dolan & Rennie (2006)	UK adolescent males in care and in young offender institutions	PCL:YV	.90
Jones et al. (2006)	Videotaped cases	PCL:YV	.91
Vincent (2002)	Canadian incarcerated young offenders	PCL:YV	.92
Skeem & Cauffman (2003)	Videotaped cases	PCL:YV	.98

In relatively few cases, researchers have reported ICCs for the individual items of the PCL. Lynam et al. (2007) report the item, facet, and total score ICCs with respect to the PCL:SV. The item scores ranged from poor (.20 for the item “poor behavioural controls”) to excellent (.86 for the item “lacks goals”), with an average item ICC of .61. The four facet scores ranged from good to excellent, with ICCs of .59, .71, .84, and .65 for the facets of Interpersonal, Affective, Behavioural, and Antisocial, respectively. Finally, the overall total score reliability was excellent, at .86. Skeem and Cauffman (2003) likewise analyzed interrater reliability for the total, factor, and item scores of the PCL:YV. Towards the conclusion of their training, six researchers rated videotaped cases

using the PCL:YV. Using a two-way mixed effects model, Skeem and Cauffman (2003) identified that the total score interrater reliability was nearly perfect ($ICC = .98$), while scores for Factor 2 were also excellent ($ICC = .95$), and good for Factor 1 ($ICC = .75$). Skeem and Cauffman then employed a weighted k to determine the interrater reliability of the 20 individual PCL:YV items. Interrater reliability was determined to range from poor (weighted $k = .20$, item = impulsivity) to excellent (weighted $k = .80$, item = impersonal sexual behaviour), with an average fair reliability of .49.

METHOD

Interrater reliability is an essential component of determining the utility of an instrument. Without confirming that raters are reliable in their measurement of a phenomena, it cannot be presumed that measures of a phenomena are reliable, which subsequently leads to an inability to determine if an instrument is valid. The CAPP has recently been developed by Cooke, Hart, and Logan (unpublished manuscript) as a comprehensive assessment tool for psychopathic personality disorder. The CAPP is currently being employed in a British Columbia youth custody research project; however, its utility as a measure of psychopathy has yet to be determined. Before the CAPP can be accepted as a useful measure of psychopathic personality disorder, it must be determined whether it is a reliable and valid measure of the construct of interest. This thesis explores this first component, in assessing the interrater reliability of the CAPP in a sample of Canadian incarcerated serious and violent young offenders.

Participants

The Serious and Violent Incarcerated Young Offenders Study is an ongoing SSHRC¹ study with incarcerated young offenders in British Columbia, Canada. Youth between the ages of 12 and 17 who are charged and/or convicted of criminal offences in Canada are either held in remand or sentenced to open or secure custody in youth custody facilities, three of which are located in British Columbia. In 2003, the Canadian government passed the *Youth Criminal Justice Act*, which stipulated that youth who are

¹ Three successive Social Sciences and Humanities Research Council Grants awarded to Raymond Corrado, Principal Investigator, in addition to a Master's level SSHRC awarded to the author

first-time or non-violent offenders should be diverted out of the traditional criminal justice system. As a result, youth in British Columbia, who are held in remand and/or incarcerated for a crime, are typically held in custody for repeat and/or violent offending.

The Serious and Violent Incarcerated Young Offenders Study is a research study in which university-level research assistants interview young offenders who are held in custody. The purpose of this study, the current version of which has been operating since 2005, is to collect diverse self-report and official information on the risk factors associated with criminality. Information is collected from the youth's court and corrections files, while, in addition, the youth provide self-report information in the context of a one-on-one confidential interview. Permission to conduct this study in British Columbia Youth Custody Facilities was obtained from the Assistant Deputy Minister of the British Columbia Ministry of Children and Families. In addition, ethics approval was obtained from the Ethics Review Board at Simon Fraser University.

All youth incarcerated in two British Columbia youth custody facilities are approached by an interviewer who invites them to participate in a university-based interview. The youth are informed that the purpose of the study is to collect information on the possible risk factors related to criminal offending. Youth are informed that there are no physical risks to participating in the interview, although the interview may touch on uncomfortable topics such as physical or sexual abuse. The youth are reassured that they are under no obligation to participate, and that if they do not wish to answer particular questions during the interview, then they may choose not to. Youth are also informed that they can end the interview at any point. Furthermore, youth are reassured that the information they share with the interviewer will be held confidential, with the

exception of a direct threat made against themselves or someone else. Youth who agree to participate in the interview sign an informed consent page, and are thanked for their participation with the provision of a pop and a bag of chips to be consumed during the interview.

The current study involved the participation of 30 incarcerated young male and female offenders. The demographic information was missing for one youth; however, for the rest of the youth, the overwhelming majority were male (96.6%), and the mean age was 15.72 years ($SD = 1.22$), ranging from a minimum of 12 years to a maximum of 18 (see Table 5). Youth custody laws generally apply to youth who are between the ages of 12 and 17; however, there are several ways in which 18 year olds may be found in Canadian youth custody settings. A youth who commits a crime between the ages of 12 and 17 may not be remanded into custody until after their 18th birthday. However, given that the youth committed their crime prior to becoming an adult, they will be remanded into youth custody. Any additional crimes that are committed at 18 years or older should subsequently be dealt with in the adult court system. It is also possible that youth who are sentenced to time in custody remain in youth custody past their 18th birthday, depending on the length of their sentence. Therefore, a youth who is incarcerated at 17 for 9 months may turn 18 years of age while still in custody; these youth will serve the remainder of their time in the youth custody facility.

The demographic information for the sample is shown in Table 5. Most of these youth were Caucasian (69%), followed by Aboriginal (20.7%), East Indian/South Asian (6.9%), or Asiatic (3.9%). At the time of their current offence, approximately one-third (34.5%) of the youth were primarily living with their natural mother. Slightly more than

half (55.2%) were enrolled in school at the time of the offence. The youth's mean age of first contact with the police was 11.86 years ($SD = 2.69$), ranging from age five to age fifteen.

Table 5: Demographic Information of Sample

	n	Percent
<i>Gender</i>		
Male	28	96.6
Female	1	3.4
<i>Ethnicity</i>		
Caucasian	20	69.0
Aboriginal	6	20.7
East Indian/South Asian	2	6.9
Asiatic	1	3.4
<i>Living with at time of offence</i>		
Natural mother	10	34.5
Foster Parents	4	13.8
Friend / Boyfriend / Girlfriend	4	13.8
Both Natural Parents	3	10.3
Natural Father	3	10.3
Group Home	2	6.9
Alone	1	3.4
With a Relative	1	3.4
On the Street	1	3.4

Information on past charges was available for 28 of the 30 youth. The youths' total number of past charges ranged from 0 to 89, with an average number of 11.93 ($SD = 36.28$). This variable was highly skewed, with the median number of charges falling at 12. The 89 charges for the one youth were primarily due to charges of breach of conditions (61% of the 89 charges), but this youth was also versatile in their criminal behaviours, having been charged in the past with multiple counts of possession of stolen

property under \$5,000, as well as several driving offences, assault charges, motor vehicle theft, robbery, and drug charges.

The most common past charge for the whole sample was theft under \$5,000, characterizing 50% of the sample (see Table 6). Slightly less than half had past charges of failing to comply with a sentence (46%), breach of conditions (39%), break and enter (36%), or possession of stolen property under \$5,000 (36%). One-quarter of the sample had past charges of robbery, breach of an undertaking, motor vehicle theft, or a driving offence. Further, one-quarter of the sample had past “other” charges, including, for example, careless use of a firearm, wilful obstruction, and carrying a concealed weapon.

Table 6: Past Charges for Sample of Serious and Violent Incarcerated Youth (n = 28)

Charge	n	Percent
Theft under \$5,000	14	50.0
Failure to Comply with Sentence	13	46.4
Breach of Conditions	11	39.3
Break and Enter	10	35.7
Possession of Stolen Property under \$5,000	10	35.7
Mischief under \$5,000	8	28.6
Breach of an Undertaking to Appear	7	25.0
Driving Offence	7	25.0
Motor Vehicle Theft	7	25.0
Other	7	25.0
Assault with a Weapon	6	21.4
Assault	5	17.9
Possession of Stolen Property over \$5,000	5	17.9
Theft over \$5,000	5	17.9
Flight from a Peace Officer	4	14.3
Assault causing Bodily Harm	3	10.7
Failure to Appear	3	10.7
Possession of a Break and Enter Instrument	3	10.7
Threats or Intimidation	3	10.7
Aggravated Assault	2	7.1
Drug Offence (not specified)	2	7.1
Possession of a Controlled Substance	2	7.1
Possession of a Prohibited Weapon	2	7.1
Assault of a Peace Officer	1	3.6
Attempted Motor Vehicle Theft	1	3.6
Drunk in a Public Place	1	3.6
Failure to Comply with YCJA Disposition	1	3.6
Mischief over \$5,000	1	3.6
Possession of a Dangerous Weapon	1	3.6
Suspension of a Conditional Discharge	1	3.6
Vandalism	1	3.6

Youth were primarily currently incarcerated for a violent offence, including either assault with a weapon, assault causing bodily harm, aggravated assault, assault, assault of police officer, robbery/attempted robbery, or threats/intimidation (see Table 7). Although over half (53.3%) of the youth were incarcerated for a violent offence, the single most common charge that resulted in their current incarceration was a probation violation or breach (72.4%).

Table 7: Most Serious Charge Resulting in Current Incarceration for Sample

Current Charge	N	Percent
Probation Violation / Breach	21	72.4
Assault with a Weapon	9	31.0
Motor Vehicle Theft	6	20.7
Driving Offence	5	17.2
Theft/Attempted Theft	5	17.2
Assault	5	17.2
Robbery / Attempted Robbery	4	13.8
Threats / Intimidation	4	13.8
Mischief	4	13.8
Possession of Stolen Property	4	13.8
Assault Causing Bodily Harm	3	10.3
Break and Enter / Attempted Break and Enter	3	10.3
Attempted Motor Vehicle Theft	3	10.3
Aggravated Assault	2	6.7
Assault of a Police Officer	1	3.3
Possession of Drugs	1	3.3

Procedure

All youth who are approached by an interviewer are offered participation in two main interviews. Prior to the interviews occurring, the interviewer conducts a review of the youth's file information. The purpose of the file review is to collect collateral information that can be used as a basis for further insight into the youth's functioning. For instance, the files may contain information provided by those who are close to or relatively familiar with the youth, such as their parents, their teachers, or their probation officers. These informants can often provide invaluable information as to the typical functioning of the youth across time, during different situations and contexts, and with a wide range of people. Information provided by these sources may reveal, for example, whether a youth tends to manipulate only their peers, or whether they have also attempted to manipulate their teachers, parents, and/or siblings. This information may also suggest the extent of time this trait has been present and whether it appears to reflect a relatively

stable trait. In addition, the collection of such “official” data will allow the interviewer to suspect when a youth is not being completely truthful. The collection of collateral information can suggest to the interviewer potential areas for further exploration.

The primary interview involves the administration of a structured questionnaire that collects information on a wide range of risk factors for youth crime, including substance abuse history, family problem profile, educational attainments, and the youth’s mental health profile. The primary interview can typically be completed in one session, lasting approximately 60 to 90 minutes. This interview is entirely closed-ended, requiring that the youth simply select the most appropriate answer from those answers provided on the questionnaire.

The second interview is semi-structured, involving an exploration of the youth’s personality profile. During this interview, youth are asked general questions about their functioning with respect to their attachment, behavioural, cognitive, dominance, emotional, and self-styles. If the youth’s response indicates that impairment in functioning might be present, interviewers are instructed to follow up with additional questions. This method allows for interviewers to be guided as to the totality of questions to be covered, yet allows them the freedom to explore potential impairments in detail.

This second interview, which collects information to be used in the CAPP evaluation, focuses on the youth’s personality styles and traits. Following the completion of this session, which can range anywhere from one to four or five sessions (approximately one hour per session) depending on the youth’s level of comprehension and willingness to talk openly, the interviewer completes an evaluation of the youth’s relative personality strengths and weaknesses.

Interviewers for the current study consisted of undergraduate ($n = 3$) and graduate ($n = 3$) students in criminology. Four interviewers received a one-day training session on the instrument by one of the authors of the CAPP (Dr. Stephen Hart), while the additional two interviewers were trained by the current author (Amanda McCormick).

For the purposes of the current study, the CAPP interviews were conducted with pairs of interviewers. Six interviewers were paired into three teams, and each interviewer-pair conducted interviews with 10 youth. In all, 30 youth were interviewed, 10 by each pair. Interviewer-pairs remained the same through the interrater reliability assessment.

The interviews were conducted with both interviewers present; during the first interview, one interviewer would ask the first half of the questions, while the second interviewer would ask the last half of the questions. In the second interview, these positions would be reversed. This method was repeated to allow each interviewer to become familiar with the nuances of each section of questions. Following the conclusion of this interview, each interviewer independently completed a CAPP evaluation. The interviewers were aware that they were being assessed for interrater reliability, but they were instructed not to discuss any case with each other, either before, during, or after their assessment of the youth. Therefore, all 30 youth were evaluated by two interviewers, all who completed the evaluation independently of their pair.

Instrument

As measured on the CAPP, traits or symptoms such as *Antagonism* and *Domineering* are assessed along a seven-point scale, where zero (not present) represents the absence of the dysfunction, and six (very severe) represents the maximum degree of severity of the dysfunction. In determining the appropriate level of identified dysfunction,

the rater must consider the persistence of the symptom (item) in question over a range of times, situations, and relationships. In order for a symptom to be assessed as “very mild”, the rater would need evidence that, while the symptom appears to be present, there is no evidence that it has resulted in dysfunction, impairment, or distress in the person’s relationships. In contrast, to assess a symptom as “very severe”, the rater would need evidence that the symptom has been present for a significant amount of time, and is persistent across contexts (e.g. it is manifested in both social situations as well as the workplace), and across relationships (e.g. it is present whether the individual is engaging with significant others, family, or peer groups). The ratings must be anchored in evidence in order to avoid confirmatory bias, where the rater seeks out information that confirms their preliminary impression of the individual (Cooke, Hart, & Logan, unpublished manuscript).

It therefore is entirely possible for an individual to display moderate severity regarding all the six symptoms reflective of the dominance domain, whereas another individual could manifest extreme interpersonal dysfunction with respect to two or three symptoms reflective of a dominant interpersonal style. This assessment technique is consistent with the idea that personality pathology is a continuous phenomenon, in which individuals may display varying degrees of dysfunction.

In addition to rating the presence of dysfunctional symptoms along a continuous scale, the CAPP also provides for dimensional assessments of illustrative indicators along a scale of zero to three. In this case, a score of zero represents that the indicator is “not at all descriptive” while a score of three represents “very descriptive”. Therefore, when rating the *antagonistic* symptom of the dominance domain, an evaluation regarding the

extent to which each of the illustrative indicators of *hostile*, *disagreeable*, and *contemptuous* effectively describes an individual's personality is completed.

The CAPP also provides for a dichotomous assessment of the presence or absence of behaviours supportive of the symptom. For instance, when assessing the degree of dysfunction relative to the symptom *antagonistic*, the rater evaluates whether behavioural indicators such as “argues with others for no good reason” or “demeans or criticizes others” are present. This comprehensive method of assessing each single symptom of personality, theoretically, allows for a more detailed overall description of personality dysfunction and enhances our understanding of the nature of PPD. Furthermore, the hierarchical nature of the model allows for the formation of heterogeneous subgroups, resulting from the varying combinations of dimensions and the degree to which their dysfunctional symptoms are present.

The CAPP is designed for application in multiple settings such as forensic or clinical settings, as well as in community or family settings. The particular version to be used in this study is the CAPP-IRS, or Institutional Rating Scale. The CAPP is a never-before used instrument that has the potential to broaden our understanding of personality dysfunction. However, before its use can be accepted in clinical settings, research is necessary to determine whether the CAPP is a reliable and valid assessment tool. The current study completes the first part of this necessary research, by documenting the interrater reliability of the CAPP.

Statistical Analysis

Interrater reliability involves an assessment of whether two or more raters are assigning similar scores to the same target, or in the current case, individual. Interrater reliability is measured using intraclass correlation coefficient (ICC) analysis. To produce ICCs, the ratio of between-groups variance to total variance is analyzed, similar to an analysis of variance (ANOVA) procedure. Essentially, ICCs are a measure of the amount of variance that may be attributed to the target, or object of measurement (McGraw & Wong, 1996; Shrout & Fleiss, 1979). With respect to the values of the ICC, as variations in the ratings decrease, the ICC will reach 1.0, or perfect reliability. Establishing high levels of interrater reliability allows one to determine that the source of any variation is the target of measurement itself. To determine the extent to which this is true requires that multiple measures of randomly selected targets must be analyzed (McGraw & Wong, 1996).

The strength of the ICCs can be determined using the following guidelines presented by Cicchetti, Showalter, and Tyrer (1985): less than .40 is poor, .40 to .59 is fair, .60 to .74 is good, while .75 to 1.00 is excellent (see Table 8).

Table 8: Guidelines for Assessing Interrater Reliability using Intraclass Correlation Coefficients

Lower Bound	Upper Bound	Reliability Determination
.75	1.00	<i>Excellent</i>
.60	.74	<i>Good</i>
.40	.59	<i>Fair</i>
-1.00	.39	<i>Poor</i>

There are multiple models of interrater analysis when using intraclass correlation coefficients. The model selected depends upon the nature of the raters and the rating target. The raters themselves may be defined either as a random sample of all possible raters whose results can therefore be generalized to the larger population of raters, or as non-random, whose results can therefore not be generalized to the larger population of raters. For instance, a randomly selected sample of 6 raters from a larger total population of 20 raters would allow the interrater reliability statistics of the six to be applied to the overall population of 20. However, if the original 6 raters were not chosen randomly, it is possible that these raters differ from the rest of the sample in some meaningful way that would not allow the results to be attributed to the rest of the population. Such a situation could occur if the six raters whose interrater reliability was analyzed were all trained by the same person, whereas the rest of the population of raters was not. Random sampling would increase the potential for such differences to be distributed equally across members of the selected group.

Similarly, the target rated can be a random sample of all possible targets, or can include all targets. Furthermore, the reliability analysis can be conducted either on the individual ratings or on the mean ratings of all raters. If future ratings will be conducted by only one rater, the individual ratings should be analyzed for interrater reliability, whereas if future ratings will be conducted by several raters, the mean ratings between all raters can be analyzed for similarity.

In addition, the correlations between the measures can be defined as either consistent or absolute measures of agreement, based on the extent to which the ratings are desired to be similar. Absolute definitions require raters to rate the phenomena

identically, and so will consider any differences between the measures as disagreements between raters. In contrast, consistency ratings examine the extent to which interrater assessments are relatively consistent with each other (McGraw & Wong, 1996).

The current analysis employed a two-way mixed model analysis, whereby a random sample of targets was rated by the fixed set of rater pairs. The raters were all those who had been hired to work as a research assistant in the larger Serious and Violent Incarcerated Youth research study; therefore, the rater sample was not a randomly selected subset of raters. Using this model of analysis limits the generalization of the ICCs to only the current sample of raters.

The current analysis used a consistency measure of agreement, to determine to what extent the ratings were consistently similar, as opposed to identical. Absolute measures would require that raters are consistently giving exactly the same response. However, the CAPP employs a seven-point rating scale, which will likely result in much variation in the extent to which raters determine a trait to be present. As such, it was presumed that consistency ratings of reliability would be a more appropriate reflection of interrater reliability, thereby allowing for small differences in the application of the seven-point scale, rather than requiring that all six raters be identical in their application of the scale to the youth participating in the current study.

Lastly, given that only one rater will typically complete a CAPP evaluation in future research, the current analysis reports the single measure of reliability, as opposed to the mean ratings of all raters. Mean ratings would be more appropriate if future research considered the assessment of multiple raters; however, once the interrater

reliability analysis of the current six raters was determined, raters proceeded to interview on their own, producing only a single assessment with respect to the CAPP.

ICCs were analyzed for the total score, six overall domain scores, and 33 item (symptom) scores for all six raters together and in their particular rater pairs. To compute the domain and total scores, the 33 symptom scores were summed for each case. The internal consistency reliabilities of the six domain scores were also tested, using Cronbach's alpha (α).

RESULTS

Description of CAPP Scores

The CAPP domain and total score descriptives can be seen in Table 9. With the exception of the attachment domain, which was significantly positively skewed, all CAPP domain scores were distributed evenly. The CAPP total scores ranged from a minimum of 0 to a maximum of 92, with a mean value of 47 ($SD = 23.11$). The CAPP total scores were distributed evenly in this sample.

Table 9: Descriptives of CAPP Domains and Total Scores

<i>Domains</i>	Mean	Median	SD	Skew (SE)	Kurtosis (SE)	Minimum	Maximum
Attachment	5.40	3.00	4.84	.92 (.43)	-.05 (.83)	0	18
Behavioural	11.30	11.50	5.63	.41 (.43)	1.13 (.83)	0	26
Cognitive	7.13	6.00	4.37	.74 (.43)	.29 (.83)	0	17
Dominance	7.50	6.50	4.86	.37 (.43)	-.95 (.83)	0	16
Emotional	7.50	7.00	4.92	.27 (.43)	-1.04 (.83)	0	16
Self	8.17	8.00	5.19	.42 (.43)	-.42 (.83)	0	19
Total	47.00	42.50	23.11	.09 (.43)	-.71 (.83)	0	92

To determine the consistency of the domains in measuring the construct of psychopathy, internal consistency reliability of the CAPP domains scores were assessed using Cronbach's alpha. Acceptable levels of internal consistency reliability as measured with Cronbach's alpha typically exceed .70, meaning that they explain at least 50% ($0.70 \times 0.70 = 0.49$, or 49%) of the variance of the construct of interest (Vogt, 2007). Presumably, scales with an alpha that is less than this amount of the variance are not a

useful representation of the construct of interest, as they measure less than half of its variance.

The alpha for all six domain scores as a scale measuring the construct of psychopathy was acceptable, at $\alpha = 0.87$ (Table 10). Using SPSS 14.0 reliability analysis, alphas were also computed to determine the effect on the overall CAPP scale should one of the domains be deleted. None of the alpha values changed significantly at this point, indicating that all six domains were reliably contributing towards the internal consistency of the CAPP scale, together contributing towards 76% of the variance of the overall construct.

Table 10: Internal Consistency Reliability of CAPP Total Score from Domain Scores

<i>Domains</i>	Alpha (α)	Alpha (α) if item deleted from scale
Attachment	.87	.84
Behavioural	.87	.85
Cognitive	.87	.86
Dominance	.87	.81
Emotional	.87	.82
Self	.87	.87

The internal consistency reliability of each of the six domains themselves was also obtained by analyzing the contribution of their respective items (Table 11). Overall, the strongest consistency was found for the attachment domain, which reached an internal consistency of $\alpha = 0.85$, while the weakest was emotional, at $\alpha = 0.67$, dropping below traditionally acceptable levels (Vogt, 2007).

Table 11: Internal Consistency Reliabilities of Domain Scores from Items

	Alpha (α)	Alpha (α) if item deleted from scale
<i>Attachment</i>		
Detached	.85	.78
Uncommitted	.85	.88
Unempathic	.85	.83
Uncaring	.85	.74
<i>Behavioural</i>		
Lacks Perseverance	.79	.79
Unreliable	.79	.76
Reckless	.79	.68
Restless	.79	.79
Disruptive	.79	.74
Aggressive	.79	.78
<i>Cognitive</i>		
Suspicious	.72	.69
Lacks Concentration	.72	.70
Intolerant	.72	.62
Inflexible	.72	.71
Lacks Planfulness	.72	.66
<i>Dominance</i>		
Antagonistic	.74	.66
Domineering	.74	.67
Deceitful	.74	.71
Manipulative	.74	.69
Insincere	.74	.72
Garrulous	.74	.75
<i>Emotional</i>		
Lacks Anxiety	.67	.58
Lacks Pleasure	.67	.68
Lacks Emotional Depth	.67	.49
Lacks Emotional Stability	.67	.68
Lacks Remorse	.67	.63
<i>Self</i>		
Self-Centred	.79	.75
Self-Aggrandizing	.79	.71
Sense of Uniqueness	.79	.75
Sense of Entitlement	.79	.79
Sense of Invulnerability	.79	.75
Self-Justifying	.79	.77
Unstable Self-Concept	.79	.83

Several items appeared problematic, as the analysis suggested that their removal may improve the reliability of the scale. For instance, the alpha for attachment was $\alpha = 0.85$, yet the analysis suggested that by removing the item “uncommitted”, reliability could be increased to $\alpha = 0.88$. Similarly, dropping the item “garrulous” from the domain of dominance would result in a one unit increase from $\alpha = 0.74$ to $\alpha = 0.75$. One point increases would also be realized on the emotional domain ($\alpha = 0.67$) by dropping either of the items “lacks pleasure” or “lacks emotional stability” (increased to $\alpha = 0.68$ in either case). Finally, with respect to the self domain, dropping the item “unstable self concept” would increase the internal consistency reliability from $\alpha = 0.79$ to $\alpha = 0.83$, an increase that may be worthwhile considering.

Moderator Variables

It is possible that ratings on the CAPP are affected by other variables such as gender of the youth, ethnic background, criminal offence, or age. For instance, in their meta-analysis of 21 studies using the PCL with youth, Edens et al. (2006) identified that the ability of the PCL to predict violent recidivism depended upon the ethnic composition of the sample. As the sample became more ethnically heterogeneous with a greater proportion of non-White youth, the ability of the PCL to predict violent recidivism decreased. With respect to gender, there is inconsistent evidence determining whether the PCL should be used with females. The PCL was created using a sample of incarcerated male adult offenders, and while research has suggested it can be used meaningfully with incarcerated adult youth offenders, the same has not been justified for females. Vincent, Odgers, McCormick, and Corrado (under review) examined the predictive validity of the PCL:YV among a sample of incarcerated males ($n = 201$) and females ($n = 55$) and

identified that the PCL:YV did not predict either violent or non-violent female recidivism. Similarly, with a sample of 125 incarcerated adolescent females, Odgers, Reppucci, and Moretti (2005) determined that the PCL:YV was unable to significantly predict either concurrent (relational or physical) or future (average follow up of 250 days) aggression among females. Furthermore, as discussed previously, psychopathic traits tend to predict more violence among young offenders, and therefore it is possible that scores will depend upon the nature of the current offence. It is also possible that scores fluctuate depending on the age of the youth, given potentially important factors such as greater maturity levels and ability to understand time perspectives (Edens et al., 2001) Given this past research, it is possible that variables such as gender, ethnicity, criminal offence, and age moderate the effect of ratings on the CAPP. Therefore, several analyses were performed to determine the potential effect of these variables on the CAPP total scores (see Table 12).

To determine the effect of ethnicity on CAPP total scores, a one-way analysis of variance was performed. The results suggested that ethnicity did not have a significant effect on CAPP total scores, $F(3, 25) = 1.66, p > .05$. Using a correlation analysis, age was not determined to have a significant relationship with CAPP total scores, $r = -.09, p > .05$. An independent samples t-test also determined that whether the offence was violent or non-violent did not have a significant effect on CAPP total scores, $t(28) = .87, p > .05$. Given that only one youth in the sample was female, it was not possible to determine the potential effects of their gender on the total or domain CAPP scores.

Table 12: Moderator Variable Effects on the CAPP Total Scores

Total CAPP Scores	Statistic	Significance
<i>Ethnicity</i>	$F = 1.66$	$p = .20$
<i>Age</i>	$r = -.09$	$p = .63$
<i>Current Violent Offence</i>	$t = .80$	$p = .43$

The effects of these variables were also compared to domain scores and found to not have an effect (see Table 13). A one-way ANOVA was run between ethnicity and each of the six CAPP dimensions; no significant results were found. A correlation analysis was performed between the six CAPP dimensions and age; again, no significant results were found. With the exception of the self domain, all correlations between age and the CAPP domains were negative. Finally, an independent samples t-test was performed between each of the six CAPP dimensions and current violent offence, and none of the results were found to be significant. The non-significant results suggest that ethnicity, age, and current violent offence did not have an effect on the ratings of the six CAPP dimensions. Again, the effects for gender could not be determined, given that there was only one female participant.

Table 13: Moderator Variable Effects on CAPP Domain Scores

Domain Scores	Ethnicity		Age		Violent Offence	
	<i>F</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>t</i>	<i>p</i>
Attachment	0.76	0.53	-0.02	0.91	-.27	0.79
Behavioural	1.56	0.22	-0.30	0.12	0.05	0.96
Cognitive	1.21	0.33	-0.22	0.25	0.76	0.45
Dominance	1.12	0.36	-0.03	0.86	1.38	0.18
Emotional	0.94	0.44	-0.04	0.83	0.44	0.66
Self	1.88	0.16	.19	0.33	1.49	0.15

Interrater Reliability

The ICCs were computed using SPSS version 14.0 reliability analysis, with confidence levels set to 95%. Confidence levels involve a lower and upper bound within which the sample statistics are expected to be found if the sample was drawn repetitively. Generally, the smaller the confidence level, the more confidence one can have that the results were not likely due to chance, but instead were likely due to some meaningful difference in the phenomena of interest. For example, a 95% confidence interval of .70 to .80 would suggest that a researcher can be 95% confident that with repeated drawings of the sample, the true sample statistic will fall somewhere between .70 and .80.

Domain scores were computed for each youth by summing together the individual symptom scores as assessed by each rater. For example, the attachment domain score represents the four summed symptom scores of detached, uncommitted, unempathic, and uncaring, each of which is rated on a scale of 0 (absent) to 6 (very severe). Overall, the ICCs indicated good to excellent total and domain interrater reliability (see Table 14). The total CAPP scores had an overall interrater reliability (among all pairs) of .91 (excellent), with a confidence interval of .82 to .96. Domain interrater reliability ranged from a low of .69 (good) to a high of .86 (excellent).

Table 14: Total and Domain Intraclass Correlation Coefficients for all Rater Pairs

	Intraclass Correlation Coefficient	Confidence Interval Lower Bound	Confidence Interval Upper Bound
Total Score	.91	.82	.96
Domains	.83	.78	.87
Attachment	.86	.73	.93
Behavioural	.86	.73	.93
Cognitive	.86	.72	.93
Dominance	.83	.68	.92
Emotional	.86	.72	.93
Self	.69	.44	.84

The results for the self-domain were substantially lower than for the remaining five domains. While the ICCs for the attachment through emotional domains ranged from .83 to .86 (CI = .68 to .93), the ICC for the self-domain was .69, with a wide confidence interval (CI = .44 to .84). A review of the seven self-domain symptoms indicated that the low reliability was primarily due to ICCs of .50 and below for the symptoms of self-centred (S1), sense of entitlement (S4), sense of invulnerability (S5), and unstable self-concept (S7), indicating only fair reliability (see Table 15). In addition, the lower bounds for several of the self-concept individual items fell into the “poor” range of reliability, and the confidence intervals for some self-concept items were quite broad. For instance, the confidence interval for unstable self-concept (S7) ranged from a lower bound of .13 to an upper bound of .70. The remaining three self-domain symptoms of self-aggrandizing (S2), sense of uniqueness (S3), and self-justifying (S6) all exceeded .60, indicating good interrater reliability. The only other individual item that demonstrated only fair reliability amongst all raters was the uncommitted (A2) symptom of the attachment domain. This symptom received an ICC of .56 (CI = .26 to .77).

Table 15: Item Intraclass Correlation Coefficients for all Rater Pairs

<i>Items</i>	Intraclass Correlation Coefficient	Confidence Interval Lower Bound	Confidence Interval Upper Bound
A1	.89	.78	.95
A2	.56	.26	.77
A3	.89	.78	.95
A4	.63	.36	.81
B1	.80	.62	.90
B2	.69	.44	.84
B3	.69	.44	.84
B4	.83	.67	.91
B5	.81	.63	.90
B6	.65	.38	.82
C1	.84	.68	.92
C2	.80	.62	.90
C3	.73	.50	.86
C4	.74	.53	.87
C5	.83	.67	.92
D1	.68	.43	.84
D2	.86	.72	.93
D3	.66	.40	.82
D4	.72	.49	.86
D5	.71	.48	.85
D6	.81	.64	.90
E1	.65	.38	.82
E2	.78	.59	.89
E3	.78	.60	.89
E4	.67	.42	.83
E5	.78	.58	.89
S1	.45	.12	.70
S2	.65	.38	.82
S3	.76	.56	.88
S4	.45	.11	.69
S5	.50	.18	.73
S6	.60	.31	.79
S7	.46	.13	.70

The total score, domain total, and item total ICCs for each pair of raters were also analyzed. To create total CAPP scores for each youth, the six domain scores were summed together to represent a total score. Domain scores were created through summing the individual symptom (item) scores for each youth. Item scores represent the individual symptom ratings given by each rater, i.e. the attachment domain would have

four item ratings for the symptoms of detached, uncommitted, unempathic, and uncaring. The total, domain, and item scores for each pair of raters were subsequently analyzed to determine the extent to which they were consistent between pairs. The ICCs for total and domain scores can be seen in Table 16, while the ICCs for item scores can be seen in Table 17.

Table 16: Total, Domain, and Item Intraclass Correlation Coefficients for each set of Rater Pairs

	Rater Pair 1			Rater Pair 2			Rater Pair 3		
	ICC	C.I. Lower	C.I. Upper	ICC	C.I. Lower	C.I. Upper	ICC	C.I. Lower	C.I. Upper
Total Score	.92	(.70)	(.98)	.89	(.61)	(.97)	.96	(.84)	(.99)
Domain Total	.84	(.74)	(.90)	.73	(.59)	(.83)	.96	(.94)	(.98)
<i>Attachment</i>	.97	(.88)	(.99)	.69	(.14)	(.91)	.93	(.74)	(.98)
<i>Behavioural</i>	.45	(.21)	(.83)	.89	(.62)	(.97)	.97	(.89)	(.99)
<i>Cognitive</i>	.88	(.58)	(.97)	.66	(.09)	(.90)	.99	(.94)	(.99)
<i>Dominance</i>	.92	(.72)	(.98)	.72	(.20)	(.92)	.97	(.87)	(.99)
<i>Emotional</i>	.90	(.66)	(.98)	.78	(.33)	(.94)	.95	(.81)	(.99)
<i>Self</i>	.80	(.38)	(.95)	.63	(.05)	(.89)	.95	(.82)	(.99)
Item Total	.75	(.70)	(.80)	.59	(.51)	(.66)	.94	(.92)	(.95)

Rater Pair 1 had excellent total score (.92), domain total (.84), and item total ICCs (.75). However, when examining the individual domain scores, Rater Pair 1 had only fair reliability when assessing the Behavioural domain. Further inspection indicated that two items in particular, unreliable (B2) and reckless (B3), were particularly low, with ICCs of .35 and .07 respectively, indicating that rater agreement was no better than chance (see Table 7). Two additional items, self-centred (S1) and unstable self-concept (S7) were also

poor; however, the domain reliability for Self for Rater Pair 1 remained excellent (ICC = .80).

Table 17: Item Intraclass Correlation Coefficients for each set of Rater Pairs

Items	Rater Pair 1			Rater Pair 2			Rater Pair 3		
	ICC	C.I. Lower	C.I. Upper	ICC	C.I. Lower	C.I. Upper	ICC	C.I. Lower	C.I. Upper
A1	.97	(.90)	(.99)	.76	(.30)	(.94)	.96	(.84)	(.99)
A2	.*	-	-	-.07	(-.64)	(.60)	.95	(.82)	(.99)
A3	.88	(.60)	(.97)	.82	(.42)	(.95)	.97	(.90)	(.99)
A4	.87	(.55)	(.97)	.11	(-.53)	(.67)	.81	(.40)	(.95)
B1	.70	(.17)	(.92)	.54	(-.14)	(.87)	.97	(.88)	(.99)
B2	.35	(-.32)	(.79)	.49	(-.16)	(.84)	.95	(.83)	(.99)
B3	.07	(-.56)	(.65)	.49	(-.16)	(.84)	.95	(.80)	(.99)
B4	.73	(.22)	(.92)	.84	(.48)	(.96)	.95	(.83)	(.99)
B5	.67	(.12)	(.91)	.87	(.57)	(.97)	.89	(.61)	(.97)
B6	.69	(.15)	(.91)	.54	(-.10)	(.86)	.97	(.89)	(.99)
C1	.90	(.62)	(.98)	.28	(-.38)	(.76)	.96	(.86)	(.99)
C2	.83	(.45)	(.95)	.71	(.18)	(.92)	.90	(.65)	(.97)
C3	.87	(.57)	(.97)	.08	(-.55)	(.65)	.94	(.79)	(.99)
C4	.91	(.69)	(.98)	.04	(-.58)	(.63)	1	(1)	(1)
C5	.67	(.07)	(.92)	.81	(.40)	(.95)	.95	(.82)	(.99)
D1	.86	(.54)	(.96)	.70	(.16)	(.96)	.95	(.80)	(.99)
D2	.89	(.63)	(.97)	.46	(-.25)	(.85)	.95	(.81)	(.99)
D3	.77	(.31)	(.94)	.37	(-.29)	(.80)	.95	(.80)	(.99)
D4	.68	(.13)	(.91)	.82	(.42)	(.95)	.86	(.54)	(.96)
D5	.72	(.21)	(.92)	.51	(-.14)	(.85)	.96	(.86)	(.99)
D6	.78	(.33)	(.94)	.75	(.28)	(.93)	.93	(.74)	(.98)
E1	.70	(.13)	(.92)	.54	(-.10)	(.86)	.89	(.63)	(.97)
E2	.36	(-.31)	(.79)	.63	(.04)	(.89)	.97	(.89)	(.99)
E3	.84	(.47)	(.96)	.69	(.11)	(.91)	.91	(.69)	(.98)
E4	.86	(.52)	(.96)	.28	(-.39)	(.76)	.96	(.86)	(.99)
E5	.87	(.56)	(.97)	.60	(-.01)	(.88)	.87	(.55)	(.97)
S1	.31	(-.36)	(.77)	-.11	(-.67)	(.53)	.95	(.80)	(.99)
S2	.72	(.21)	(.92)	-.11	(-.67)	(.53)	.93	(.75)	(.98)
S3	.89	(.61)	(.97)	.74	(.24)	(.93)	.91	(.69)	(.98)
S4	.81	(.41)	(.95)	-.28	(-.75)	(.39)	.64	(.06)	(.89)
S5	.59	(-.02)	(.88)	.29	(-.38)	(.76)	.96	(.84)	(.99)
S6	.70	(.17)	(.92)	.37	(-.30)	(.79)	.70	(.17)	(.92)
S7	-.16	(-.69)	(.49)	.58	(-.03)	(.88)	.72	(.22)	(.92)

* There was no variance for this item, as the ratings given were all 0.

The ICCs of Rater Pair 2 were not as strong as Rater Pair 1. The total score ICC remained excellent, however, the domain total and item totals dropped to good (.73) and fair (.59) respectively. Although none of Rater Pair 2's domain score reliabilities dropped

below good (i.e. all domain ICCs exceeded .60), there were some indications of problems when the 33 item ICCs were reviewed. 12 of the 33 item scores (36%) for rater pair 2 dropped below “poor” interrater reliability, which is more than what would be expected by chance alone. Specifically, the ICCs for the symptoms of uncommitted (A2), uncaring (A4), suspicious (C1), intolerant (C3), inflexible (C4), deceitful (D3), lacks emotional stability (E4), self-centred (S1), self-aggrandizing (S2), sense of entitlement (S4), sense of invulnerability (S5), and self-justifying (S6) were all less than .40. Further, an additional five item scores only reached “fair” interrater reliability. Surprisingly, given the number of self symptoms with poor reliability for Rater Pair 2, the ICC for the Self domain was good (.63); however, the confidence interval for this ICC was quite large (CI = .05 to .89).

Rater Pair 3 had the highest ICCs, with excellent reliability for the total CAPP score (.96), the domain total (.96), and the item totals (.94). Furthermore, all individual domain total scores for Rater Pair 3 were excellent, with all ICCs exceeding .93. In reviewing all of the 33 item ICCs, none of the symptom ratings were found to be of poor reliability (i.e. all item ratings exceeded .40); furthermore, only three of the individual item scores dropped below excellent reliability. For Rater Pair 3, sense of entitlement (S4), self-justifying (S5), and unstable self-concept (S7) all had ICCs in the “good” range.

Correlations between Dimensions of the CAPP

Vincent (2002) identifies that in evaluations of an instrument, the primary aspect to determine is whether the test is structurally valid, whether it is consistent with its theoretical structure. She then notes that measures of psychopathy can be defined as

unidimensional, as the items included in the instrument all pertain to one overarching construct: psychopathy. However, Vincent further notes that measuring a unidimensional construct does not prevent the presence of underlying correlated factors. Analyses of data collected with the PCL-R and PCL:YV have suggested that psychopathy can be factored into either two, three, or four factors. In all cases, these factors tend to be moderately correlated with each other, indicating that they are all related to a common construct or coherent syndrome, in this case psychopathy.

Psychopathy as proposed by Cooke, Hart, and Logan (unpublished manuscript) is a six-dimensional hierarchical model. In effect, psychopathy is a superordinate construct that is underpinned by six individual, yet related, domains of functioning. As such, it is expected that these six dimensions will be moderately correlated with each other and more strongly correlated to the construct of psychopathy (as measured by total scores) itself. To test this hypothesis, correlations were obtained between each of the six dimensions and between each dimension individually with the total score.

The results generally supported the hypothesis (see Table 18). All domains had a strong, significant relationship with total scores. While most of the domains showed moderate correlations between each other, three results did not fit this pattern. The relationship between the emotional domain and the attachment domain was particularly strong, at $r = .84$.

Table 18: Correlations between Domain and Total CAPP Scores

	Attachment	Behavioural	Cognitive	Dominance	Emotional	Self	Total
Attachment	—						
Behavioural	.52**	—					
Cognitive	.41*	.62**	—				
Dominance	.66**	.59**	.57**	—			
Emotional	.84**	.49**	.47**	.72**	—		
Self	.40*	.30	.28	.70**	.49**	—	
Total	.81**	.76**	.70**	.90**	.85**	.68**	—

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

The pattern between domains also did not hold for the self domain with respect to the behavioural domain ($r = .30$) or the cognitive domain ($r = .28$). In both cases, the relationship, while significant, was substantially lower than other relationships identified. Furthermore, the correlation between the total CAPP scores and the self domain, while strong and significant, was smaller than all other domains with respect to the total CAPP scores. Although smaller than the other domain and total score correlations, the correlations were still substantial enough to be considered strongly related to the total scores. Given the results presented in Table 18, the strongest relationship with respect to CAPP total scores appears to be the dominance domain, with a correlation of $r = .90$.

DISCUSSION

Criticisms regarding the ability of the PCL to tap into clinically relevant items of psychopathy have led others to re-conceptualize the disorder. In contrast to the PCL, which is structured by the domains of interpersonal, affective, and behavioural functioning, the CAPP provides insight into six domains of functioning: attachment, behavioural, cognitive, dominance, emotional, and the self. Theoretically, the CAPP should therefore provide greater insight into dysfunction. Furthermore, in recognizing that PPD is a hierarchical construct composed of multiple heterogeneities of symptoms, the CAPP should pave the way to a greater understanding of the construct of PPD.

Prior to examining the interrater reliability, the effects of moderator variables including ethnicity, age, and nature of the offence (violent or non-violent) was tested with respect to the total scores and domain scores of the CAPP. No significant results were found, suggesting that in the current sample of 30 youth, ethnicity, age, and violent or non-violent offence did not affect the CAPP scores. With the exception of the self domain, all correlations between age and CAPP domains were negative. It is not surprising that the self domain was positively, although not significantly, correlated with age, as one would expect the self domain to be more stable as age increases, due to a greater understanding of who one is as a person and self-appreciation for one's strengths and weaknesses.

Internal consistency reliability analyses also suggested that the CAPP total score, composed of the six domain scores, has acceptable levels of reliability. Further, the domain scores themselves generally had acceptable levels of reliability. The exception in this case was the emotional domain, which fell below the generally acceptable alpha of

.70. Examining the items composing the emotional domain revealed that the internal consistency reliability of the domain would not be substantially improved by the removal of any one symptom. It is essential that the internal consistency of this domain be explored in additional studies to determine whether the symptoms of emotional functioning are truly unrepresentative of the construct, or whether the results apply only to the current sample of 30 serious and violent incarcerated young offenders.

The results of the internal consistency reliability also suggested that the self domain scale may be improved with the removal of the symptom “unstable self-concept”, as removing this one item would improve the internal consistency of the scale. Prior to any symptom being removed from the scale, however, it is important to replicate this study in additional research to determine whether this symptom continues to be problematic in contributing towards understanding of self-styles of functioning. It is possible that this particular symptom fell below acceptable levels of internal consistency due to the nature of the sample, i.e. adolescents.

The results of the analysis indicate that the total and domain interrater reliability of the CAPP is generally good to excellent, while the item interrater reliability is substantially lower. The results of the current analysis are in line with past research that has been conducted with the PCL and PCL:YV. As previously discussed, interrater reliability analyses using intraclass correlation coefficients with measures obtained through administration of the PCL instruments typically ranges between 0.79 and 0.98 for total scores. Similarly, in the current analysis, the CAPP total score interrater reliability was excellent, with an overall average ICC = .92. The total CAPP scores for each set of

rater pairs also fell within this range; all were excellent at $ICC_1 = .92$, $ICC_2 = .89$, and $ICC_3 = .96$.

Analyses of factor scores obtained from the PCL instruments are also similar to the results obtained in the current study. Skeem and Cauffman (2003) identified ICCs of .75 and .95 for the PCL Factor 1 (Interpersonal/Affective) and Factor 2 (Behavioural), respectively. In the current analysis, average factor, or domain, scores ranged from good at .69 to excellent at .86. The paired reliability results of Rater Pair 1 did fall below this general range, with a low of .45 for the self domain. However, they achieved excellent reliability on the attachment domain, with an ICC of .97. The domain reliabilities of Rater Pair 2 also fell slightly below what Skeem and Cauffman (2003) reported, with a range of ICCs from .63 to .89. Rater Pair 3 produced substantially better interrater reliability results, with a range of ICCs from .93 to .99, thus exceeding the factor scores identified by Skeem and Cauffman (2003).

As discussed previously, it is relatively uncommon for authors to present item ICCs for interrater reliability analyses employed with the PCL. In one of the few studies that has, Lynam et al. (2007) identified a range of ICCs that were as low as .20 (poor reliability) to a high of .86 (excellent). Similarly, the current study identified a wide range of item ICCs, falling below zero and reaching as high as .99. Lynam et al. (2007) reported that the average item ICC was .61; relatively similar results were identified in this current study. The average item interrater reliability was .75 (excellent), .59 (fair), and .94 (excellent), for Rater Pairs 1, 2, and 3, respectively.

Despite consistencies with prior research, there do appear to be several areas of particular concern. Specifically, the overall interrater reliability of the self-domain was

particularly low amongst all raters, and in reviewing the domain and item scores, this domain appeared to be the most inconsistent for all individual rater pairs. The problems with this domain appeared to be the result of low interrater consistency in evaluating the symptoms of self-centred, sense of entitlement, sense of invulnerability, and unstable self-concept.

This result is clarified further by the contribution in particular of the unstable self-concept item to the measurement of the self-domain. The results of the internal consistency reliability, as measured by Cronbach's alpha, of the self domain suggested that the reliability of this scale could be improved from .79 to .83 with the removal of the item "unstable self-concept". This result implies that this item may be particularly difficult for raters to assess consistently, leading to variable measurements of the overall construct of self-functioning.

The relative inconsistency in evaluating the extent of dysfunction related to the self-concept may be a direct result of the youth's age itself. Given that youth are continuing to develop both physically and mentally, self-functioning styles may not yet be fully established. This can potentially result in fluctuations in self-concept and self-presentation that may result in difficulties in accurately assessing a youth's stable sense of self. While further training may assist in reducing the inconsistency in this domain, it is highly possible that the comparatively low interrater reliability is a function of target age, as opposed to rater inconsistency.

In reviewing the interrater reliability for each of the three rater pairs, there did not appear to be any particular items which were consistently problematic. While both Rater Pair 1 and Rater Pair 2 had poor reliability in assessing self-centred (S1), Rater Pair 3 had

excellent reliability for this symptom. Similarly, while Rater Pair 2 had several other symptom ratings that were “poor”, these symptom ratings were generally good to excellent for the other two rater pairs.

While Rater Pair 1 had difficulties in reliably assessing unreliable (B2) and reckless (B3) symptoms, the behavioural symptom ratings for Rater Pair 2 and Rater Pair 3 did not produce any ICCs below “fair” reliability. Further, it appears that the especially low rating for Rater Pair 1 with respect to “reckless” (ICC = .07) may be due to inconsistent application of a reference group. In discussions with these raters following the conclusion of the interrater reliability analysis, it was determined that one rater was evaluating “reckless” by comparing the youth in question to the general population, while the other rater was evaluating “reckless” by comparing the youth to other incarcerated youth. Youth who are charged with a criminal offence and held in a youth custody facility generally tend to be more reckless than the general population, and this inconsistent application of a reference group would therefore subsequently lead to conflicting assessments with respect to extent to which the youth in question appears to be “reckless”.

A possible explanation for the less overall consistent ratings made by Rater Pair 2 is that while both Rater Pair 1 and Rater Pair 3 conducted the primary interviews with the youth, given time constraints, Rater Pair 2 conducted only the second interview and CAPP evaluations. Although Rater Pair 2 reviewed the completed primary interview and file review for each of the 10 youth, it is possible that the lower symptom reliability scores were a function of less familiarity with the youth in question. In addition, the interviews conducted by Rater Pair 2 typically were spread across a period of several

weeks. This was due to the availability of Rater Pair 2 to interview only one day a week, and given that interviews with a single youth were typically spread across several interviewing sessions, Rater Pair 2 often took several weeks to conclude their interview and assessment. In contrast, the interviews conducted by Rater Pair 1 and Rater Pair 3 were typically conducted in the space of a week or two. Therefore, it is highly possible that the length of time between the start of the interview and the CAPP evaluation resulted in lower symptom reliability scores for Rater Pair 2. It is important to note, however, that despite the poor reliability for several symptoms, the total score reliability, domain score reliability, and total item reliability for Rater Pair 2 ranged between good to excellent. However, these results do suggest that in assessing the CAPP, it is important that raters conduct multiple interviews with the individual in question, or at least attempt to conclude the CAPP interview in a shorter period of time.

The correlations between domains, and between domain and total CAPP scores were also assessed. The model of psychopathy proposed by Cooke, Hart, and Logan (unpublished manuscript) is hierarchical, with psychopathy representing a superordinate construct composed of six individual, but related, domains. The expectation that the six domains would be moderately correlated, while more strongly correlated to the hierarchical construct of psychopathy, was generally supported. However, correlations were stronger than expected between the emotional and attachment domains of functioning ($r = .84$). This suggests a potential lack of independence between these two areas of functioning. Although it is expected that these two domains will overlap, as one presumes that capacity for emotion is inherent in the ability to form relationships with others, it suggests that these domains may not actually be measuring separate constructs.

The pattern between domains also did not hold for the self domain with respect to the behavioural domain ($r = .30$) or the cognitive domain ($r = .28$). In both cases, the significant correlation was substantially lower than expected. It is possible that perceptions of self are not strongly influenced by cognitive processes or that they do not strongly influence behavioural tendencies. Of course, it is also possible that the relationship between these domains is partially suppressed, given the relatively poor interrater reliability for the self domain. In effect, the self domain had only “good” overall reliability, and it is likely that raters had difficulty in assessing these traits, which may have led to the subsequent weaker relationship with other domains of functioning. It is possible that in future research, or in research with adults, these relationships will become stronger.

Methodological issues and limitations

The importance of collaborative information when evaluating personality structure and dysfunction cannot be overstated. People are typically motivated to present their best side to others, and offenders are certainly no exception to this general rule; in fact, offenders often have a vested interest in presenting their best side to prison psychologists or others who are assessing their progress while in prison. The tendency to present the best side of oneself highlights the importance of accessing documented file information when completing an assessment of an offender’s personality profile. The insufficient file information in a number of our cases may have impeded the interviewers’ ability to accurately evaluate the personality structure of the young offender. This phenomenon may also have led to the sometimes discrepant evaluations of particular symptoms. It is certainly plausible for one evaluator to discount the sincerity of an

offender's responses while the other embraces it wholeheartedly; this is simply a function of who we are as humans. Therefore, similar to the use of the PCL scales in clinical settings, the CAPP should always be used in conjunction with documented information and the evaluator should never base their opinion solely on information provided by the offender. While not all offenders set out to deceive and manipulate interviewers, such behaviour can unfortunately not be put past a certain percentage of them.

The CAPP provides a detailed rating scale with which an evaluator can determine the presence/absence of a trait and the degree of severity of the dysfunction. This rating scale is much more comprehensive than that used with the PCL, which simply allows for the absence, likely presence, and definite presence of a trait. However, there is something to be said for simplicity, in that a rating scale with only three points of reference is certainly likely to result in greater reliability between interviewers. The CAPP scales of severity for each symptom is measured on a seven-point scale, and slight differences in opinion between interviewers are to be expected. In the majority of discrepancies, the differences in opinion were minor; for instance, a two versus a three. However, in a small number of cases these discrepancies were substantially wider. While concerning, it must be noted that these situations were relatively infrequent, and that overall, the interrater reliability of the CAPP total scores was determined to range from good to excellent.

Prior research suggests that the psychopathy construct may not be the same in females as it appears to be in males. Research has suggested that, to date, the PCL:YV has limited utility in predicting criminal justice outcomes with females. However, in the current study, one of the youth involved in the CAPP interrater reliability was female. Given that only one youth was female, it was not possible to determine whether gender

would have a significant effect on the rating of CAPP symptoms and subsequent domain and total scores.

In addition, there are some limitations to the instrument itself. The CAPP can involve a lengthy interview and often requires several sessions with the youth before it can be completed. For Rater pair 2, at least, this may have been the source of some discrepancy in the ratings given, as Rater pair 2 were only able to complete one interview a week, resulting often in interviews taking two to three weeks before they were completed. Furthermore, given the current length of the interview, it is not possible to determine the test-retest reliability of the instrument within the current sample, as it is not expected that the average youth will have the patience to sit through the same lengthy interview twice within their time in prison. Despite these limitations, the current study was able to provide the first results stemming from the use of the CAPP with incarcerated populations, and provides the first step in determining its appropriateness for use.

When beginning to use the CAPP in clinical and incarcerated settings, it is recommended that interviewers follow the procedures used in the current study. Given the extent of information collected and analyzed in conducting a CAPP evaluation, it is essential that interviewers are initially paired. Furthermore, although interviewers involved in the reliability analysis were not permitted to debrief each other following the interview evaluation, it is recommended that future interviewers debrief each other consistently following interviews until one can be sure that interviewers are interpreting information to the same degree. When there is a lack of corroborating information for use with the CAPP interview and evaluation, another recommendation would be to simply caution against the use of the CAPP for clinical purposes, given the potential for negative

legal outcomes, such as using an evaluation as an aggravating factor in sentencing or other criminal justice decision making. Alternatively, two interviewers could evaluate the same individual and their average ratings could be taken as reflective of the offender's personality structure. A final recommendation on this topic is to encourage evaluators to tape record their interviews and to document descriptively the logic behind their scoring decisions. When one is able to explain the process by which they came to their conclusion, greater confidence can be had in the results.

Future Research

The initial results provided by this study lend support to the interrater reliability of the CAPP for total and domain scores. Individual item scores, however, were characterized by less stable ratings. Future research must continue to explore the nature of interrater reliability in additional settings. The current analysis involved a two-way mixed model of interrater reliability, whereby a random sample of targets were rated by the fixed set of rater pairs, thereby rendering the results non-generalizable to other populations. Studies that use the CAPP in future research, whether in similar settings or not, should continue to assess the interrater reliability of this instrument. The CAPP total and domain score interrater reliability is well within the range of prior research; however, future research should further explore whether particular items or domains continue to persist with respect to inconsistent rater evaluations.

Given that the current analysis was performed using data collected from serious and violent incarcerated youth in a Canadian setting, it is imperative that other studies further explore the role of the self domain in additional populations, whether adult custody, forensic psychiatric settings, non-Canadian institutions, or community samples.

It is possible that the comparatively low interrater reliability for this domain was the result of the relatively young age and maturity level of the sample. The sample was characterized by young, relatively impulsive youth, who are likely still exploring their identity, trying on different roles while they attempt to define their self-concept. As such, it would not be surprising that research with adult samples finds stronger interrater reliability for this domain.

Additional validations of this pilot instrument must be empirically tested in order to support its utility in criminal justice and clinical settings. The current analysis provided the first steps towards determining the CAPPs usefulness in a prison setting; however, the analysis was limited to an examination of the reliability of the instrument. Specifically, in future research the CAPP must be evaluated with respect to its construct, criterion, and predictive validity.

An instrument is regarded as valid if evidence supports that it accurately measures what it intends to measure. In analyzing whether the CAPP demonstrates validity, future research should test the ability of the CAPP to correlate with measures of similar constructs, such as the Psychopathy Checklist, and to discriminate from measures of non-similar constructs. In essence, validity analyses should test the ability of the CAPP to measure the theorized traits included in its content.

In addition, the utility of an instrument such as the CAPP is its ability to predict the likelihood of certain outcomes. The PCL has demonstrated its predictive validity numerous times. For instance, Corrado et al. (2004) supported the predictive validity of the PCL-YV with respect to general and violent re-offending of young offenders released from custody. Therefore, future research should endeavour to assess the utility of the

CAPP in predicting future behaviours such as general and violent criminality. In summary, the validity of the CAPP has yet to be demonstrated, and until such time, the CAPP should be restricted for use in research settings.

CONCLUSION

The results of this analysis suggest that the CAPP has utility as a measure of psychopathy, in that it produces acceptable levels of internal consistency reliability and interrater reliability. The CAPP provides a comprehensive assessment of personality functioning, and while developed as a measure of psychopathic personality disorder, it may also provide insight into the relative strengths and weaknesses of those who are assessed with it.

Measures obtained using the CAPP will allow us to broaden our understanding of the personality traits that are associated with delinquency. In the current study, youth were rated on the CAPP following a review of their file information and after participating in two sets of interviews with a research assistant. Their subsequent scores, as determined by the rater, should provide insight into their particular areas of vulnerability. For instance, a youth who is incarcerated for a violent crime may have issues relating to several of these domains. They may, in particular, have problems committing to relationships with others, and allowing themselves to depend on other people. They may also be suspicious of others, and as a result, attempt to dominate their relationships. Using the CAPP with large numbers of incarcerated youth will in the future reveal whether patterns of personality are present in relation to particular forms of offending. While the PCL has allowed us to determine, for example, that psychopathic individuals tend to commit more crime, both general and violent, and at a faster rate than non-psychopaths, it has not necessarily broadened our understanding of the causes of these behaviours. In addition, the PCL's utility in determining the causes of criminality has been inhibited by its reliance on antisocial behaviours such as juvenile delinquency or

criminal versatility. Including such measures in an assessment of psychopathy, while improving the predictive validity of the measure, inhibit our ability to comprehend the causes of behaviours such as juvenile delinquency or criminal versatility.

Further, by assessing a six-dimensional model of personality with reference to 33 symptoms of functioning, it should be possible to determine particular areas of vulnerability to where treatment may be targeted. Therefore, the CAPP is not limited to use with those suspected of having psychopathic personality disorder. Its utility should apply more broadly to offenders in general, and ideally in future research it will assist in broadening our understanding of the causes and correlates of criminality.

Once validated as a useful measure of psychopathy, the CAPP will also offer utility in longitudinal research. The CAPP can be applied as a measure of lifetime functioning, or it can be used in terms of more discrete periods of functioning. For instance, an offender may be assessed using the CAPP upon entry at a custodial facility. At this time, questions may refer to consistent, lifetime patterns of functioning. In order to determine whether treatment offered during the course of incarceration has had an effect, prior to release from custody the offender can again be assessed with the CAPP, with the questions relating specifically to the period of functioning over the last, for example, six months to one year. Differences in rating scores may then provide insight into whether there appears to have been any treatment success.

The CAPP provides a window into the mind of the psychopath. The ability to determine personality dysfunction using a six-dimensional, hierarchical model allows for the recognition of heterogeneity in this disorder. Using a six-dimensional model of psychopathy will provide opportunity in future research to explore possible subtypes of

offenders, both psychopathic and not. Improving our understanding of the heterogeneity and subsequent subtypes of this disorder will allow us to further comprehend the aetiological processes that underlie the development of psychopathic personality disorder. Essentially, personality is reflected in behaviour, and so if we are ever to develop a better understanding of the aetiology of criminal behaviours, we must first improve understanding of personality. The CAPP provides such a tool, enabling a more comprehensive understanding of personality dysfunction.

The current study supported the interrater reliability of the CAPP. Future research should endeavour to explore the validity of this instrument. If found to be both a reliable and valid assessment of psychopathy, the CAPP will have exceptional utility in criminal justice and mental health settings. Furthermore, given that the CAPP provides for a more in depth exploration of the symptomatology behind personality dysfunction, it will enable a greater understanding of the true nature of psychopathy. Ideally then, the CAPPs utility will also be applicable to treatment settings. While no known treatment has yet been identified for psychopathic personality disorder, perhaps the reason is not that psychopathic individuals are untreatable, but that they are treated as a homogeneous group. Given the CAPPs ability to differentiate personality dysfunction through the identification of 33 symptoms across six broad domains, the CAPP may provide future researchers and clinicians with the ability to tease apart subgroups of offenders, and to better tailor treatment options and criminal justice responses accordingly.

REFERENCES

- Barry, C.T., Frick, P.J., & Killian, A.L. (2003). The relation of narcissism and self-esteem to conduct problems in children: A preliminary investigation. *Journal of Clinical Child and Adolescent Psychology, 32*: 139-152.
- Blackburn, R., Logan, C., Renwick, S.J.D., Donnelly, J.P. (2005). Higher-order dimensions of personality disorder: Hierarchical structure and relationships with the five-factor model, the interpersonal circle, and psychopathy. *Journal of Personality Disorders, 19*(6): 597-623.
- Blair, R.J.R. (2006). Subcortical brain systems in psychopathy: The amygdala and associated structures. In C.J. Patrick (ed.), *Handbook of Psychopathy* (pp.296-312). New York: Guilford Press.
- Blair, K.S., Newman, C., Mitchell, D.G.V., Richell, R.A., Leonard, A., Morton, J., & Blair, R.J.R. (2006). Differentiating among prefrontal substrates in psychopathy: Neuropsychological test findings. *Neuropsychology, 20*(2): 153-165.
- Blonigen, D.M., Carlson, S.R., Krueger, R.F., & Patrick, C.J. (2003). A twin study of self-reported psychopathic personality traits. *Personality and Individual Differences, 35*: 179-197.
- Burns, G.L. (2000). Problem of item overlap between the Psychopathy Screening Device and Attention Deficit Hyperactivity Disorder, Oppositional Defiant Disorder, and Conduct Disorder rating scales. *Psychological Assessment, 12*(4): 447-450.
- Cicchetti, D.V., Showalter, D., & Tyrer, P.J. (1985). The effect of number of rating scale categories on levels of interrater reliability: A monte carlo investigation. *Applied Psychological Measurement, 9*(1): 31-36.

- Cleckley, H. (1964). *The Mask of Sanity*. (4th ed.). St Louis, MO: Mosby.
- Cooke, D.J., Hart, S.D., & Logan, C. (unpublished manuscript). *The Comprehensive Assessment of Psychopathic Personality Disorder*.
- Cooke, D.J. and Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment, 13*(2): 171-188.
- Cooke, D.J., Michie, C., Hart, S.D., & Clark, D.A. (2004). Reconstructing psychopathy: Clarifying the significance of antisocial and socially deviant behavior in the diagnosis of psychopathic personality disorder. *Journal of Personality Disorders, 18*(4): 337-357.
- Corrado, R.R., Vincent, G.M., Hart, S.D., and Cohen, I.M. (2004). Predictive validity of the Psychopathy Checklist: Youth Version: For general and violent recidivism. *Behavioral Sciences and the Law, 22*, 5-22.
- Crawford, T.N., Cohen, P., & Brook, J.S. (2001). Dramatic-erratic personality disorder symptoms: I. Continuity from early adolescence into adulthood. *Journal of Personality Disorders, 15*(4): 319-335.
- Dolan, M. & Blackburn, R. (2006). Interpersonal factors as predictors of disciplinary infractions in incarcerated personality disordered offenders. *Personality and Individual Differences, 40*: 897-907.
- Dolan, M. & Rennie, C. (2006). Psychopathy Checklist: Youth Version and Youth Psychopathic trait Inventory: A comparison study. *Personality and Individual Differences, 41*: 779-789.

- Edens, J.F., Campbell, J.S., & Weir, J.M. (2006). Youth psychopathy and criminal recidivism: A meta-analysis of the Psychopathy Checklist measures. *Law and Human Behavior, 31*: 53-75.
- 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 and the Law, 19*: 53-80.
- Everitt, B.S. (1996). *Making Sense of Statistics in Psychology: A Second Level Course*. New York: Oxford University Press.
- Farrington, D.P. (2005). The importance of child and adolescent psychopathy. *Journal of Abnormal Child Psychology, 33*(4): 489-497.
- Frick, P.J. & Hare, R.D. (2001). *The Antisocial Process Screening Device*. Toronto, ON: Multi-Health Systems.
- Forth, A.E., Kosson, D.S., & Hare, R.D. (2003). *The Psychopathy Checklist: Youth Version*. Toronto, Ontario: Multi-Health Systems.
- Gray, J. (1987). *The Psychology of Fear and Stress*. Cambridge: Cambridge University Press.
- Gray, J. & McNaughton, N. (1995). The neuropsychology of anxiety: reprise, [in:] Nebraska symposium on motivation. In D. Hope, & C. Izard (eds.), *Perspectives on Anxiety, Panic, and Fear* (pp.61-134). Lincoln: University of Nebraska Press.
- Greenberg, J.R. & Mitchell, S.A. (1983). *Object Relations in Psychoanalytic Theory*. Cambridge, MA: Harvard University Press. Referenced in Benjamin, L.S. (1996). An interpersonal theory of personality disorders. In J.F. Clarkin & M.F.

- Lenzenweger (eds.), *Major Theories of Personality Disorder* (pp 141-220). New York: Guilford Press.
- Gretton, H.M., Hare, R.D., & Catchpole, R.E.H. (2004). Psychopathy and offending from adolescence to adulthood: A 10-year follow-up. *Journal of Consulting and Clinical Psychology, 72*(4): 636-645.
- Hall, J.R. & Benning, S.D. (2006). The “successful” psychopath: Adaptive and subclinical manifestations of psychopathy in the general population. In C.J. Patrick (ed.), *Handbook of Psychopathy* (pp.459-478). New York: Guilford Press.
- Hare, R.D. (2003). *The Hare PCL-R* (2nd Edition). Toronto, Ontario: Multi Health Systems.
- Hare, R.D. (1998). Psychopathy, affect and behavior. In D.J. Cooke, A.E. Forth, and R.D. Hare (eds), *Psychopathy: Theory, Research, and Implications for Society* (pp. 105-137). Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Hare, R.D. (1993). *Without Conscience: The Disturbing World of Psychpaths Among Us*. New York: Guilford Press.
- Hare, R.D. (1991). *The Hare Psychopathy Checklist-Revised (PCL-R)*. Toronto, Ontario: Multi-Health Systems.
- Hare, R.D., Forth, A.E., & Strachan, K.E. (1992). Psychopathy and crime across the life span. In R.D.V. Peters, R.J. McMahon, & V.L. Quinsey (eds.), *Aggression and Violence Throughout the Life Span* (pp. 285-300). Newbury Park, CA: Sage Publications.

- Hare, R.D. & Neumann, C.S. (2006). The PCL-R assessment of psychopathy: Development, structural properties, and new directions. In C.J. Patrick (ed.), *Handbook of Psychopathy* (pp.58-88). New York: Guilford Press.
- Harris, G.T., Rice, M.E., & Quinsey, V.L. (1994). Psychopathy as a taxon: Evidence that psychopaths are a discrete class. *Journal of Consulting and Clinical Psychology*, 62(2): 387-397.
- Hart, S.D, Cox, D.N., & Hare, R.D. (1995). *The Hare Psychopathy Checklist: Screening Version*. Toronto, ON: Multi-Health Systems.
- Hart, S., Watt, K., & Vincent, G.M. (2002). Commentary on Seagrave and Grisso: Impressions of the state of the art. *Law and Human Behavior*, 26: 241-245.
- Jackson, W. & Verberg, N. (2007). *Methods: Doing Social Research* (4th ed.). Toronto: Pearson Prentice Hall.
- Jones, S., Cauffman, E., Miller, J.D., & Mulvey, E. (2006). Investigating different factor structures of the Psychopathy Checklist: Youth Version: Confirmatory factor analytic findings. *Psychological Assessment*, 18(1): 33-48.
- Karpman, B. (1941). On the need for separating psychopathy into two distinct clinical types: Symptomatic and idiopathic. *Journal of Criminology and Psychopathology*, 3, 112-137.
- 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 and the Law*, 21: 771-786.
- Lilienfeld, S.O. (1998). Methodological advances and developments in the assessment of psychopathy. *Behaviour Research and Therapy*, 36: 99-125.

- Lilienfeld, S.O. (1994). Conceptual problems in the assessment of psychopathy. *Clinical Psychology Review, 14*(1): 17-38.
- Lilienfeld, S.O. & Andrews, B.P. (1996). Development and preliminary validation of a self-report measure of psychopathic personality traits in noncriminal populations. *Journal of Personality Assessment, 66*(3): 488-524.
- Lykken, D.T. (1995). *The Antisocial Personalities*. New Jersey: Lawrence Erlbaum Associates.
- Lykken, D. (1957). A study of anxiety in the sociopathic personality. *Journal of Abnormal Psychology, 55*: 6-10.
- Lynam, D.R., Caspi, A., Moffitt, T.E., Loeber, R., & Stouthamer-Loeber, M. (2007). Longitudinal evidence that psychopathy scores in early adolescence predict adult psychopathy. *Journal of Abnormal Psychology, 116*(1): 155-165.
- Lynam, D.R., Caspi, A., Moffitt, T.E., Raine, A., Loeber, R., & Stouthamer-Loeber, M. (2005). Adolescent psychopathy and the big five: Results from two samples. *Journal of Abnormal Child Psychology, 33*(4): 431-443.
- Lynam, D.R., & Derefinko, K.J. (2006). Psychopathy and personality. In C.J. Patrick (ed.), *Handbook of Psychopathy* (pp.133-155). New York: Guilford Press.
- Marcus, D.K., John, S.L., & Edens, J.F. (2004). A taxometric analysis of psychopathic personality. *Journal of Abnormal Psychology, 113*(4): 626-635.
- Maxfield, M.G., & Babbie, E. (2005). *Research Methods for Criminal Justice and Criminology* (4th ed.). Belmont, CA: Thomson Learning, Inc.
- McGraw, K.O. & Wong, S.P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1): 30-46.

- Newman, J.P., MacCoon, D.G., Vaughn, L.J., & Sadeh, N. (2005). Validating a distinction between primary and secondary psychopathy with measures of Gray's BIS and BAS constructs. *Journal of Abnormal Psychology, 114*(2): 319-323.
- 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 and the Law, 23*: 743-763.
- Ogloff, J.R.P. (2006). Psychopathy/antisocial personality disorder conundrum. *Australian and New Zealand Journal of Psychiatry, 40*(6): 519-528.
- Porter, S. (1996). Without conscience or without active conscience? The etiology of psychopathy revisited. *Aggression and Violent Behavior, 1*: 179-189.
- Poythress, N.G. & Skeem, J.L. (2006). Disaggregating psychopathy: Where and how to look for subtypes. In C. Patrick, *Handbook of Psychopathy* (pp. 172-192). New York: Guilford Press.
- Raine, A. & Yang, Y. (2006). The neuroanatomical bases of psychopathy: A review of brain imaging findings. In C.J. Patrick (ed.), *Handbook of Psychopathy* (pp.278-295). New York: Guilford Press.
- Rutter, M. (2005). Commentary: What is the meaning and utility of the psychopathy concept? *Journal of Abnormal Child Psychology, 33*(4): 499-503.
- Salekin, R.T. (2006). Psychopathy in children and adolescents: Key issues in conceptualization and assessment. In C.J. Patrick, *Handbook of Psychopathy* (pp. 389-414). New York: Guilford Press.

- Salekin, R.T., Brannen, D.N., Zalot, A.A., Leistico, A., & 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.
- Seagrave, D., & Grisso, T. (2002). Adolescent development and the measurement of juvenile psychopathy. *Law and Human Behavior, 26*: 219-239.
- Shedler, J. & Westen, D. (2004). Dimensions of personality pathology: An alternative to the Five-Factor Model. *American Journal of Psychiatry, 161*: 1743-1754.
- Shrout, P.E. & Fleiss, J.L. (1979). Intraclass correlations: Uses in assessing reliability. *Psychological Bulletin, 86*(2): 420-428.
- 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 and the Law, 21*: 737-770.
- Skilling, T.A., Harris, G.T., Rice, M.E., & Quinsey, V.L. (2002). Identifying persistently antisocial offenders using the Hare Psychopathy Checklist and DSM antisocial personality disorder criteria. *Psychological Assessment, 14*: 27-38.
- Stouthamer-Loeber, M. (1986). Lying as a problem behavior in children: A review. *Clinical Psychology Review, 6*: 267-289.
- Ullrich, S. & Marneros, A. (2004). Dimensions of personality disorders in offenders. *Criminal Behaviour and Mental Health, 14*: 202-213.
- Vasey, M.W., Kotov, R., Frick, P.J., & Loney, B.R. (2005). The latent structure of psychopathy in youth: A taxometric investigation. *Journal of Abnormal Child Psychology, 33*(4): 411-429.

- Viding, E., Blair, R.J.R., Moffitt, T.E., & Plomin, R. (2005). Evidence for substantial genetic risk for psychopathy in 7-year olds. *Journal of Child Psychology and Psychiatry*, 46(6): 592-597.
- Vien, A. & Beech, A.R. (2006). Psychopathy: Theory, measurement, and treatment. *Trauma, Violence, & Abuse*, 7(3): 155-174.
- Vincent, G.M. (2002). Investigating the legitimacy of adolescent psychopathy assessments: Contributions of item response theory. *Unpublished Doctoral Dissertation*.
- Vincent, G.M., Odgers, C.L., McCormick, A.V., & Corrado, R.R. (Under review). Submitted to the *International Journal of Psychiatry and Law*.
- Vincent, G.M., Vitacco, M.J., Grisso, T., & Corrado, R.R. (2003). Subtypes of adolescent offenders: Affective traits and antisocial behavior patterns. *Behavioral Sciences & the Law*, 21(6), 695-712.
- Vogt, W.P. (2007). *Quantitative Research Methods for Professionals*. Boston, MA: Pearson Education, Inc.