A META-ANALYSIS OF THE ASSOCIATION BETWEEN PSYCHOPATHY CHECKLIST AND RISK ASSESSMENT INSTRUMENT SCORES

_____

A Thesis

Presented to

The Faculty of the Department of Psychology and Philosophy

Sam Houston State University

_____

In Partial Fulfillment

of the Requirements for the Degree of

Master of Arts

_____

by

Stephanie L. Goodwin

December, 2022

A META-ANALYSIS OF THE ASSOCIATION BETWEEN PSYCHOPATHY CHECKLIST AND RISK ASSESSMENT INSTRUMENT SCORES

by

Stephanie L. Goodwin

_____

APPROVED:

Marcus Boccaccini, PhD
Committee Chair

Laura Drislane, PhD
Committee Co-Chair

Jared Ruchensky, PhD
Committee Member

Chien-Pin Li, PhD
Dean, College of Humanities and Social Sciences

DEDICATION

I would like to dedicate the completion of my thesis to my family. Mum and Dad; you have always supported me, no matter how many twists and turns my path took. You always made me feel capable of anything I set my mind to. Michael, no matter how different we are and how much we disagreed growing up, I know you always have my back.

ABSTRACT

Goodwin, Stephanie L., *A meta-analysis of the association between psychopathy checklist and risk assessment instrument scores*. Master of Arts (Clinical Psychology), December, 2022, Sam Houston State University, Huntsville, Texas.

Assessing risk for future violence or reoffense is a common and important task for forensic evaluators. Indeed, these assessments are among the most common requests received by forensic evaluators. To conduct such evaluations most accurately and efficiently, it is important for forensic evaluators to have knowledge about the tools they utilize in assessments and how they may interact with and/or overlap with one another, an area of research that is greatly underdeveloped.

The current study aimed the examine the relationship between Psychopathy Checklist (PCL) measure total, factor, and facet scores and risk assessment total scores. Although the PCL family of measures were not created as risk assessment measures, they have come to be used in risk assessments due to moderate correlations between PCL scores and recidivism. Random-effects meta-analytic procedures were utilized to determine the mean correlation between PCL measures and risk assessment measures found in existent literature or received from authors of papers in the risk assessment literature. Overall, results suggest a moderate to large correlation between PCL Factor 2, facet 3, and facet 4 scores and risk assessment measures and small to moderate correlations between PCL Factor 1, facet 1, and facet 2 scores. Additionally, correlations between PCL Factor 2, facet 3, and facet 4 scores meet or exceed $r = .70$ for many specific risk assessments suggesting these PCL components may be completely redundant with preexisting risk assessment measures.

KEY WORDS: Psychopathy checklist, PCL, Risk assessment, Violence risk, Sexual Risk, Recidivism, Correlation, Meta-analysis

ACKNOWLEDGEMENTS

First, I would like to extend my appreciation to my committee for their assistance throughout this project. Marc, for assuring me that I was prepared for every hurdle I faced along the way and calming my nerves when I began to feel overwhelmed. Laura and Jared for the countless articles and books that you sent me for reference and hours your spent reviewing analyses with me and discussing my findings. Without you all, this project would not be the quality that it is, nor would I have the degree of knowledge regarding meta-analyses and the importance of this area of research.

Second, I would like to extend my appreciation to my fellow doctoral students who assisted me at various points throughout this project. Although there are too many of you who helped me with the emotional turmoil that is graduate school, there are a few that I would like to thank in particular. Rebekah, thank you for the moral and research support that you provided throughout this project. It helped that we were working on similar projects so that we could share information, worries, and successes along the way. Seeing you working and succeeding was uplifting and encouraging and your kind words and assistance did not go unnoticed. Kelci and Beata, thank you for providing a listing ear when it felt like this project would never end, or there was what appeared to be an insurmountable hurdle. Thank you also for allowing my thesis to interrupt our game nights, movie nights, and pool days, which often featured me pulling out a laptop for "just a few moments."

To my amazing friend Kelsi, thank you for all the moral support direct from Canada. Thank you for looking over portions of this project along the way for edits and allowing discussion of psychopathy and risk assessment to dominate many discussions

for months at a time. You have always been such an amazing support system for me, and I cannot wait to celebrate you and return the favor in the coming years.

Finally, to Kyle. You are my rock, my biggest support. Thank you for always somehow knowing exactly what I need in a moment and asking if there is anything you could do to ease some of the burden. I know it is not always easy, especially when I am stressed, busy, and maybe a bit hangry, but you handled it all with grace and support. You kept me on task when I was ready to give up, and made tiny steps forward seem like giant's leaps. I love you.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

CHAPTER I

Introduction

Assessing risk for future violence or reoffense is a common task for forensic evaluators. The practice of conducting risk assessments has high stakes for both offenders and community members with community safety being the priority of these assessments, and the freedom of the offender limited should they be found at high risk to reoffend. Survey findings indicate that referrals for violence risk assessments are the second most common referrals received by forensic evaluators (Neal & Grisso, 2014). Researchers have developed more than 400 measures specifically designed to assess risk for violence or reoffense (Neal & Grisso, 2014; Singh et al, 2014) and many evaluators use one or more of these measures in each risk assessment case (Neal & Grisso, 2014). Most of these measures were developed over the past 25-years, after repeated findings of poor predictive validity for opinions based in unstructured clinical judgment (Dawes, Faust, & Meehl, 1989; Grove et al, 2000; Melton et al, 2018; Monahan, 1981). Several large meta-analytic reviews have supported the conclusions that risk assessment instrument scores outperform unstructured clinical judgment and predict violence and recidivism at a level that is significantly better than chance, although there is no clear evidence, in any specific risk assessment area, that one measure consistently outperforms all others (Abbiati et al, 2019; Campbell, French, & Gendreau, 2009; Scalora, Viñas-Racionero, & Cawood, 2020; Smid et al, 2014; Yang, Wong, & Coid, 2010).

Many forensic evaluators also use Hare's Psychopathy Checklist-Revised (PCL-R; Hare, 2003) in risk assessment evaluations, usually as part of a battery that includes risk specific measures (Boccaccini et al., 2017; Jackson & Hess, 2007; Viljoen et al.,

2010). Although not specifically designed as a risk assessment measure, the PCL-R has come to be used in risk assessment due to moderate-sized correlations between PCL-R scores and future violence (Leistico et al., 2008). Existing studies provide little information about how evaluators combine or integrate findings when they use both psychopathy and risk-specific measures, or the extent to which they consider the possible overlap between psychopathic traits and risk instrument scores as part of this process. Moreover, although there are many risk assessment studies that report results for both PCL-R and risk assessment instrument scores, correlations between scores on these measures are often not reported or, when reported, are rarely discussed at length.

Because psychopathy measure scores are regularly used inform risk assessment judgments, the extent to which they are redundant with or provide information beyond risk assessment instrument scores information should be an important consideration. Thus, the current meta-analysis seeks to examine the overlap between psychopathy checklist (PCL) scores and scores on developed risk assessment measures in order to determine the extent to which existing risk assessment scores already assess psychopathic traits.

Risk Assessment Instrument Use

Most evaluators rely on risk assessment measures to help assess an individual's level of risk. Findings from one recent survey suggest that between 89.0 and 96.9% of evaluators use at least one formal risk assessment measure when coming to their conclusions about offender or patient risk (Neal & Grisso, 2014). Forensic evaluators have a large number of options in terms of risk assessment measures and each of these measures has its own, sometimes extensive, body of research. International evaluator

surveys and reviews of the risk assessment literature suggest that there are at least 400 different instruments used in violence risk assessment cases (Neal & Grisso, 2014; Singh et al, 2014). The proliferation of different instruments makes it difficult to compare across studies or applications, because each instrument assesses risk in a slightly different manner and evaluators use different batteries of assessments to conduct their assessments.

Findings from one recent survey indicated that the number of assessment measures evaluators used per evaluation ranged from 1 to 15, with an average of four per case (Neal & Grisso, 2014). Evaluators in this survey used 110 different tools in a variety of different combinations. In some studies, instruments were almost entirely redundant (*r* = .80-.88; see for example Douglas et al, 2005; Glover et al, 2017; Summer & Loza, 2004; Warren et al, 2005), which suggests combining these particular measures would not add incremental value above the score from a single measure. The issue of integrating results from multiple risk instruments is complex and generally underdeveloped in the risk assessment literature. Ideally, there would be incremental validity studies for common instrument combinations showing that scores on one measure provide information above and beyond another. Although this information exists for some measures (e.g., the STABLE-2007 and the Static-99R, see Hogan & Sribney, 2019, and the STABLE-2007 and the Risk Matrix 2000, see Helmus et al, 2015), it does not exist for many possible score combinations. In the absence of detailed incremental validity findings, it would be useful to at least know the extent to which scores on two measures are correlated. These correlations should help evaluators understand whether instruments assess redundant or unique information and if they should or should not expect similar findings across measures. For example, evaluators should expect similar findings from

two highly correlated measures because the measures probably measure similar traits or constructs (Cronbach & Meehl, 1955).

Much evidence exists to suggest that structured risk assessment measures, including structured professional judgment and actuarial instruments, outperform unstructured professional judgment in the prediction of recidivism[1] (e.g., Grove & Meehl, 2000; Povli, 1999). Studies suggest that unstructured professional judgement has low predictive accuracy for non-violent (AUC = .50), sexual (AUC = .52-.59), and violent (AUC = .52-.57) recidivism (Bengtson & Langstrom, 2004; Langton, 2003; Mori, 2017). Structured risk assessment measures, by contrast, have demonstrated moderate predictive accuracy for all types of recidivism (AUC = .56-.76; e.g., Bengtson & Langstrom, 2004; Langton, 2003; Mori, 2017).

In contrast, comparing across different structured risk assessment measures, current research suggests that structured risk assessment measures perform similarly in the prediction of recidivism. Specifically, studies have found that risk assessment measures tend to have moderate levels of predictive efficacy for any recidivism (AUC = .64-.68; Desmarais, Johnson, & Singh, 2016; Langton, 2003), violent recidivism (AUC =

[1] There are three main types of decision-making process in risk assessment; unstructured professional judgment, mechanical judgment (i.e., based on actuarial measures), and structured professional judgment. Unstructured professional judgment refers to subjective judgments made by professionals in the field of risk assessment based on interview data/knowledge of the offender, and clinical experience. Mechanical judgment, on the other hand, does not allow for any clinical interpretation of instrument results. Mechanical judgment is based on objective statistical measures that generally rely on the coding of the presence of static or historical factors that are related to recidivism. Structured professional judgment represents a middle-ground between unstructured and mechanical judgment, with the evaluator collecting and examining data concerning specific research-supported static and historical risk factors but allowing for clinical judgment in the process of using those factors to come to conclusions about future risk.

.59-.76; Bengtson & Langstrom, 2007; Langton, 2003; Mori, 2017; Smid et al, 2014), and sexual recidivism (AUC = .61-.77; Bengtson & Langstrom, 2007; Langton, 2003; Mori, 2017; Smid et al, 2014), as well as misconduct in the prison system (AUC = .65-.83; Abbiati et al, 2019). Studies directly comparing multiple risk assessment instruments suggest that no one structured risk assessment measure performs better than others for violent and sexual recidivism with minor exceptions (e.g., Barbaree et al, 2001; Langton, 2003; Yang, Wong, & Coid, 2010).

Psychopathy and the PCL Family of Measures

One individual characteristic of particular relevance to risk of violence or reoffending is psychopathy; a personality disorder defined by a pattern of interpersonal, affective, and behavioral characteristics (Hare, 2003). The dominant instrument used to assess psychopathy in correctional and forensic settings is Hare's Psychopathy Checklist-Revised (PCL-R; Hare, 1991, 2003). The PCL-R was developed as a method to measure Cleckley's criteria for psychopathy and to provide insight into important clinical presentations among prisoners. The PCL-R can be modeled using a two factor structure Factor 1 (interpersonal/affective) which consists of eight items, and Factor 2 (lifestyle/antisocial) which consists of 10 items. Each of these factors is subdivided into two facets: facet 1 (interpersonal) consisting of four items, facet 2 (affective) consisting of four items, facet 3 (lifestyle) consisting of five items, and facet 4 (antisocial) consisting of five items (Hare, 2003). Although the PCL-R was not developed as a risk assessment measure, it is nonetheless commonly used by forensic evaluators in such contexts.

The PCL-R came to be used in risk assessment because of research documenting moderate associations between PCL-R scores, violence, and recidivism (Edens & Campbell, 2007; Hawes, Boccaccini, & Murrie, 2013; Leistico et al, 2008; Mokros, Vohs, & Habermeyer, 2014), and it was available and had an established research base years before the development of many risk-specific measures. According to Hare (1999), psychopathy functions as a general predictor of both sexual and violent recidivism, however, its strength as a predictor of sexual recidivism is strongly reliant on the addition of phallometric evidence of deviant sexual arousal. A survey of forensic evaluators found that 75.6% of practitioners believed the consideration of psychopathic traits to be essential to a sex offender risk assessment, and that 65.8% of these evaluators frequently or always utilize the PCL-R as a measure of psychopathy in these cases (Jackson & Hess, 2007). Additional surveys of forensic evaluators corroborate these findings and expand upon them, finding that both the PCL-R and the PCL:SV are within the top 10 most utilized risk assessment measures in violence and sexual risk assessment evaluations as well as in evaluations for civil commitment which require a standard of risk of harm to self or the community to be met (Archer et al, 2006; Neal & Grisso, 2014; Singh et al, 2014; Viljoen. McLachlan, & Vincent, 2010). Specifically, one study found that in cases where the PCL-R was employed, it was commonly used to provide information about the risk for sexual reoffending (67.4%), the risk for violent offending (40.0%), a measure of mental illness or abnormality (45.3%), or a combination of recidivism and mental illness (31.6%; Boccaccini et al, 2017). Similarly, the PCL-YV has been found to be utilized in risk assessments of child and adolescent offenders, but with less frequency than in adult evaluations (Archer et al, 2006;) and is within the top 10 most utilized risk assessment

measures in and adolescent risk assessments (Viljoen. McLachlan, & Vincent, 2010). Forensic evaluators typically rely on PCL-R total scores, reporting factor and facet scores less than half of the time (Boccaccini et al, 2017).

PCL-R proponents, including Hare, argue that PCL-R scores should be used in conjunction with other information and/or measures (e.g., phallometric testing results or alongside standard actuarial risk assessments with the actuarial measures entered first into the prediction model), as opposed to it being a stand-alone risk assessment measure (Hare, 2003; Hemphill & Hare, 2004). The recommended practice of using the PCL-R along with other standard risk assessment measures raises questions about the extent to which PCL-R results offer something unique to risk assessments that is not already captured by scores from instruments that were designed with the express purpose of evaluating risk. The PCL-R gathers information on static risk factors related to previous criminal behavior similar to those captured by other risk assessment instruments including early behavioral problems, juvenile delinquency, revocation of conditional release, and criminal versatility. The PCL-R also overlaps with clinical aspects of risk assessments gathering information on risk factors including impulsivity, irresponsibility, lack of realistic long-term goals, promiscuous sexual behavior and parasitic lifestyle (i.e., dependence on others and avoidance of employment).

The PCL-R does, however, take into consideration information that is different from other measures. Specifically, the PCL-R gathers information on an individual's affective and interpersonal features including grandiosity, superficial charm, pathological lying, manipulativeness, and lack of empathy, remorse, or guilt. Thus, the underlying assumption in utilizing the PCL-R above and beyond currently available risk assessment

measures is that PCL scales identify offenders that are particularly sadistic, violent, and unlikely to be rehabilitated and thus will be especially predictive of criminal behavior, clinical treatability, and overall risk level due to lack of conscience.

There are now more than 200 studies that report scores for the PCL and at least one risk measure and there are more than 60 different risk measures used across studies (see Table 1). Although many of these studies do not report correlations between risk instrument and PCL scores, those that do show that correlations between the PCL and risk measure scores vary significantly, from weak (r = .18 - .39) to strong (r = .69-.77; see for example Glover et al, 2002; Hilton et al, 2008; Parent et al, 2018) and in some cases, when compared to measures that include protective factors (i.e., the SAPROF), the correlations are negative (r = -.47 - -.55; see for example De Page et al, 2018; De Vogel et al, 2019). Additionally, correlations between PCL scores and those from the same risk measure also vary depending on the study (for example r = .25-.85 for the HCR-20; see Desmarais et al, 2012; Douglas et al, 2005).

Studies that report correlations between PCL factor scores and risk measures suggest that Factor 2 scores correlate more strongly with risk measures than Factor 1 scores (r = .33-.45 for Factor 1 and r = .55-.81 for Factor 2; see for example Neal et al, 2015; Simourd & Hoge, 2000; Warren et al, 2005), and those studies that reported on correlations with facet scores indicate that facet 3 (r = .49-.71)and facet 4 scores (r = .61-.86; i.e., the lifestyle and antisocial facets) tend to correlate more strongly with risk measures than facet 1 (r = .24-.39) and facet 2 scores (r = .14-.57; i.e., the interpersonal and affective facets; see for example Churcher et al, 2016; Neves et al, 2011; Stockdale et al, 2014) although there is also significant variability across studies.

Validity of PCL-R Scores for Risk Assessment

The existing literature on the PCL-R suggests that although the total score is a significant predictor of general, sexual, and violent recidivism their predictive capabilities are weak to moderate in effect (r = .03-.49, r = -.12-.14 and r = -.02-.47 respectively; see Barbaree et al, 2001; Caperton, 2005; Larsen et al, 2020;; Sjostedt & Langstrom, 2002; Walters, 2003). Additionally, PCL-R total scores are predictive of misconduct, aggression, and institutional infractions with a small to medium effect sizes (Edens & Campbell, 2007, Larsen et al, 2020, Leistico et al, 2008). Caution into these findings has been suggested, as the association between PCL-R total scores and recidivism or institutional aggression of any type appears to depend on the sample utilized (i.e., type or severity of offending) and length of follow-up period. A meta-analysis by Hawes and colleagues (2013) suggests that there is a large variability in the predictive effects of the PCL-R in the prediction of sexual recidivism (d=.40, range -.18 to .96) and some studies suggest that PCL-R total scores are an inadequate predictor of any type of recidivism (Harris et al, 2013; Hill et al, 2012).

Unlike the total score, only 47.9% of practitioners report the two factors scores, and 30.9% report facet scores. Approximately 20.0% of evaluators consider factor 2 the most predictive of sexual recidivism and 7.4% of evaluators report facet 4 to be the most predictive of sexual recidivism (Boccaccini et al, 2017). Despite the overwhelming popularity in the belief that the total score of the PCL-R is the most predictive component of the measure, this minority may be most accurate in their assumptions. The current literature suggests that factor 2 and facet 4 scores of the PCL-R boast the highest correlations for violent and sexual recidivism out of the possible PCL-R scores. Indeed,

factor 1 scores, do not appear to be correlated with recidivism, and thus the majority of the correlation between PCL-R total scores and recidivism stem from the role of factor 2 (in particular facet 4) scores in prediction (Barbaree et al, 2001; Hawes et al, 2013; Langevin & Curnoe, 2011; Larsen et al, 2020; Sjostedt & Langstrom, 2002; Sturup et al, 2016; Walter, 2003). Sturup and colleagues (2016) even go so far as to say that facet 4 scores are the most reliable and have the best predictive validity in long-term offenders and Murrie and colleagues (2012) determined that although facet 1 characteristics may be the most unique to psychopathic offenders, they are the least predictive of recidivism. A meta-analysis by Walters (2003) suggests that 72-85% of the time, Factor 2 scores are more highly correlated with recidivism than are Factor 1 scores.

Despite evidence that Factor 2 and facet 4 scores are the strongest predictors of sexual and violent recidivism, studies suggest that Factor 1 traits are more highly associated with clinician opinions about risk. In fact, PCL-R Factor 1 scores explain over three times the variance in clinician risk judgments than do Factor 2 scores, and there appears to be no association between facet 4 scores and clinician risk judgments (Gardner et al, 2018). These correlations may occur because Factor 1 items are more prototypically associated with psychopathic offenders whereas Factor 2 items are more routine and are observed in offenders who are not labeled as psychopathic. As such, more weight may be erroneously placed on the items that seem to identify particularly sadistic and/or violent offenders who are assumed to be incapable of being rehabilitated. Additionally, these factors are more vivid and may be easier to recall after the long process of scoring the PCL-R and may be more likely to affect final risk judgments for this reason.

Although it appears that the PCL-R total scores, as well as its factor and facet scores, have moderate predictive abilities as a risk assessment instrument, there are many other instruments that have been developed with the specific purpose of aiding to predict recidivism. Current literature on the use of the PCL-R in forensic evaluations suggests the PCL-R correlates with a variety of other sex offender and violence risk assessment measures. These measures include both actuarial measures including the Static-99 family of tests, Minnesota Sex Offender Screening Tool-Revised (MnSOST-R), Violence Risk Appraisal Guide (VRAG), and Sex Offender Risk Appraisal Guide (SORAG) and structured professional judgment measures including the Historical, Clinical and Risk Management-20 (HCR-20), Sexual Violence Risk-20 (SVR-20), and Risk Matrix 2000 (RM2000; Barbaree et al, 2001; Caperton, 2005; Hill et al, 2012; Parent et al, 2011). In fact, some of these measures, including the SORAG, VRAG, SVR-20, and HCR-20, include PCL-R scores in their calculations to take psychopathy into account. In studies comparing the PCL-R to these other measures, the PCL-R total score has been found to have a lower predictive validity compared to many of the tools developed expressly for the purpose of assessing the risk of recidivism, including those that did not take psychopathy into account (Barbaree et al, 2001; Singh et al, 2011; Stadtland et al, 2005; Yang et al, 2010). The lower predictive validity of PCL-R total scores may be attributable in part to the including of content that does not predict recidivism risk (F1), which diminishes the impact of the features of psychopathy that are relevant to risk prediction (F2).

Because psychopathy has been considered both relevant, and important to risk assessment evaluation, early measures designed for the assessment of general, violent,

and sexual risk assessment, including the HCR-20, SORAG, VRAG, Domestic Risk Appraisal Guide (DVRAG), and SVR-20 included PCL-R total or factor scores in their scoring systems (Hare & Neumann, 2009). Because of the inclusion of the PCL-R in other risk assessment instruments, correlational research between PCL-R scores and scores from other instruments is complicated. A moderate correlation should be expected due to the inclusion of the same information, however, it is also important to consider the individual contributions of the additional information within these risk assessment measures.

Current Study

The existing literature on the use of the PCL-R in forensic risk assessment evaluations raises several unanswered questions about PCL-R scores in risk assessment contexts. One such question is whether PCL-R scores offer something unique to risk assessments, or if PCL-R scores are generally redundant with the information gained using other instruments that were designed with the express purpose of evaluating risk of general, violent, and sexual recidivism post-release?

One way to address this question is to examine the degree of overlap in information provided by facets of the PCL-R and other risk assessment measures. For the current study, I reviewed existing studies on psychopathy utilization in risk assessment for information about the correlation between PCL family of measures (PCL-R, PCL:SV, and PCL-YV), and other actuarial and structured professional judgment risk assessment measures. A preliminary review of the psychopathy research identified 252 articles that reported results from a PCL measure and at least one other risk assessment measures. Of these studies, 74 (29%) reported at least one correlation between a PCL score and risk

assessment instrument score, whereas 178 (70.63%) reported no correlation with a risk assessment measure. Studies also varied on which correlations they reported with 68 (27%) of the studies reporting correlations between the risk assessment measure and the PCL total score, 32 (13%) reporting correlations with PCL factor scores, and 14 (6%) reporting correlations with PCL facet scores.

Those studies that reported correlations suggested a wide range of correlations from .13 to .85 between the PCL-R and a variety of risk assessment measures including the Rapid Risk Assessment for Sexual Offense Recidivism (RRASOR), SORAG, VRAG, Static-99, HCR-20, and SVR-20 (Barbaree et al, 2001; de Vogel, Bruggeman & Lancel, 2019; Douglas, Yeomans, & Boer, 2005; Hill et al, 2012; Neal et al, 2015).

Even less research exists on the correlations between risk assessment measures utilized in adolescent risk assessments. Two studies suggested that correlations between PCL:YV and other adolescent risk assessment measures including the Structured Assessment of Violence Risk in Youth (SAVRY) and Youth Level of Service/Case Management Inventory (YLS/CMI) range from .48 to .77, and that the correlation between PCL:YV scores and Static-99 scores is approximately .35 (Viljoen et al, 2009; Welsh et al, 2008). Considering the large variety in overall correlations between the PCL family of measures and other risk assessment measures, I expect a meta-analysis on the intercorrelations of risk assessment measures to reveal moderate to large correlations between the PCL family of measures and other risk assessment measures for general, violence, and sexual recidivism, meaning the PCL measures index similar or redundant content with risk-specific measures.

Even less previous research exists on the correlations between PCL factor and facet scores and scores on other risk assessment measures. Studies that have investigated the correlations between PCL factor scores and risk measures have found correlations between .10-.56 for Factor 1 and between .51-.88 for Factor 2 (Douglas et al, 2005; Neal et al, 2015). Studies that have investigated the correlations between PCL-R facet scores and the risk measures have found correlations between .168-.46 for Facet 1, .29-.39 for Facet 2, .45-.62 for Facet 3 and .45-.50 for Facet 4 (de Vogel et al, 2019; Neal et al, 2015). De Vogel and colleagues also found negative correlations between -.196 and -.467 for each of the PCL-R Facet scores and the Structured Assessment of Protective Factors for violence risk (SAPROF). It is important to note that only two studies reviewed reported correlations between PCL-R facet and factor scores and other risk assessment measures. If we suspect that the PCL-R is redundant with risk assessment measure for the prediction of risk, it is important to know which aspect or aspects of psychopathy is contributing to this overlap. Based on this limited information, as well as what we know about the correlations between PCL-R factor and facet scores and recidivism, it is likely PCL-R Factor 2 scores that would have larger correlations with other risk assessment measures. It is also important to consider if both facets of Factor 2 are contributing to this overlap, or is the antisocial behavior and static factors (facet 4) contribute more than the disinhibited personality traits (facet 3). If this is the case, does the lack of overlap between Factor 1, facet 1, and facet 2 scores and other risk assessment measure suggest that these items are unreliable measures of criminal recidivism and violence, or do these capture aspects of criminal recidivism that are distinct from those measured by other risk assessment measures?

The goal of this study was to conduct a comprehensive review and meta-analysis of the correlation between PCL measure total, factor, and facet scores and total scores on risk assessment measures designed to estimate violent, sexual, and general recidivism. To ensure my study was comprehensive, I conducted a thorough literature search to compile a list of studies that utilized a PCL family measure (i.e., the PCL-R, the PCL:SV, or the PCL:YV) and at least one other risk assessment measure. I contacted the authors of these studies and asked them to provide additional statistical information about their studies including the correlations between measure scores, the main effect size utilized in this study. I intended to use the data collected to answer several questions about the use of PCL measures in risk evaluations. First, is there substantial overlap between PCL total, factor, and facet scores and commonly utilized risk assessment measures? Correlational data allow us to consider the amount of overlap in the measures assessed and provide information about the potential unique contribution of the PCL measures. To expand upon this area of research I was also interested in examining the degree of overlap between PCL total, factor, and facet scores and types of risk assessment measures (i.e., sexual risk assessment, violence risk assessment, domestic violence risk, protective factor measures, and general recidivism measures). This would help gather information concerning whether the PCL may provide a unique contribution in specific types of risk assessments over other types of risk assessments. These findings will be applicable to evaluators who routinely conduct risk assessment evaluations. Results can be used to determine the weight with which PCL total, facet, and factor scores should be given within a risk assessment evaluation, as well as identify any potential redundancy between information gleaned from the PCL and other risk assessment measures.

CHAPTER II

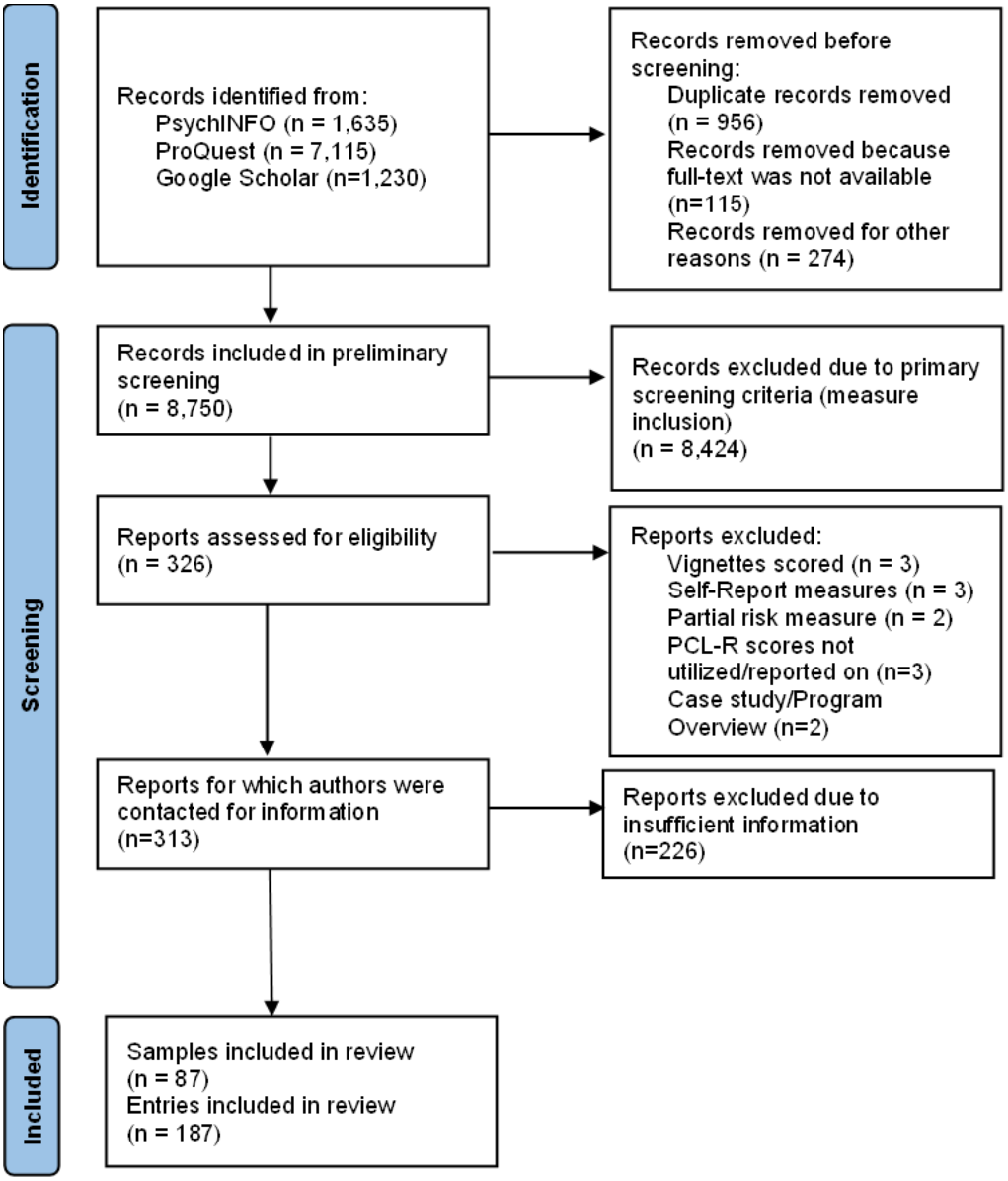
Methods

Search Strategy

I located studies for this meta-analysis by searching PsycINFO, ProQuest, and Google Scholar databases using the search terms “psychopathy” and “risk” I used filters to exclude articles written before the publication of the PCL-R measure in 1991. Additionally, only research articles available in English were screened due to language fluency restrictions. I searched for studies until January 2022 so that I would be able to include all studies published through the end of 2021. After accounting for redundant articles (i.e., available on multiple platforms), those that were not empirical studies, and those for which full-text versions could not be obtained this search process yielded 8,750 results. I screened the full-text article, thesis, or dissertation when the study abstract implied the inclusion of a PCL measure to ascertain whether the article should be considered for eligibility for my study. After screening a total of 326 articles were selected for consideration for eligibility (Figure 1).

Figure 1

PRISMA Flow Diagram for Student Inclusion in the Current Review



Inclusion Criteria

Due to the nature of the study, the examination of the correlation between risk assessment measures and the PCL family of measures, I only reviewed studies that included data pertaining to a PCL measure and at least one other risk assessment measure. Thus, I excluded studies if they did not present data on a PCL measure or on at

least one risk assessment measure despite mentioning them throughout the article. Additionally, I excluded studies if the measures were not administered to participants for clinical or research purposes. For example, several studies utilized the scoring of risk assessment measures on vignettes rather than individual participants. Because I was interested in examining the overlap among measures commonly scored by clinicians during risk assessment evaluations, I excluded studies that only used self-report measures of violence or risk. Further, I excluded studies utilizing or reporting on only a portion of a risk assessment measure, except when separate identifiable subscales were utilized to measure different types of recidivism.

Overall, three studies were excluded because the PCL and risk assessment measures were scored using vignettes, three studies were excluded because the additional risk assessment measure was a self-report measure (i.e., SAQ and the RST-i), two studies were excluded because they utilized only individual items or a subset of items from risk assessment measures, one study was excluded because the PCL-R was mentioned in the introduction, but was not utilized in the methodology, one study was excluded because the PCL-R was used to help score other measures but scores on the PCL-R were not examined independently in the results. Additionally, one study was excluded because the participants were noted to have PCL-R scores over 30, however the PCL-R was not utilized in the study, one study was excluded because it was a summary of case studies rather than an experimental design, and one study was excluded because it was a summary of different programs, but no findings were reported from the measures.

Following exclusions, 313 eligible articles remained. For samples that contained multiple additional risk assessment measures, each risk assessment measure was

considered for eligibility. In one case, a study contained both clinician administered and self-report risk assessment measures. For this study, the clinician administered and scored measures were included in the study, and the self-report measure (the SAQ) was excluded due to improper fit with the experimental question.

Article Coding

After scanning articles for inclusion fit, each article was coded for study criteria. Coding of all articles was completed by the principal investigator (S.G.).

Study Information

This section coded information on the study title and author. Additionally, information on the source of the data (e.g., journal, thesis, dissertation), year of study, and country of study was recorded.

Moderating Variables

Information regarding potential moderating variables including the type of population utilized (i.e., correctional, institutional, outpatient, etc.), demographic information (sex, age, race/ethnicity measured as percentage of sample Caucasian, country of origin), and the type of risk measure utilized (i.e., violence risk, sexual risk, general risk) was coded. If a study used both male and female participants, information was recorded about the number of participants of each sex (i.e., number of male participants and number of female participants). Due to low numbers of female participants in mixed sample and the prevalence of single sex samples, moderation analyses were conducting using only single sex samples. Similarly, if the sample included both adults and juveniles, or samples from different populations (i.e., correctional,

institutional, outpatient, etc), information about the proportion of data reflecting each same type was coded.

Measure Information

The second section of the coding form contained information regarding the PCL and risk measures. First, information was recorded about the version of the PCL utilized in the study, the method of scoring (i.e., file only or file and interview), coder training in the measure, and descriptive statistics for the total, factor, and facet scores. Second, similar information was coded for each risk assessment utilized including measure name, descriptive statistics, reason for coding (i.e., clinical or research purposes) and coder training. For each PCL and risk assessment measure reported, information was recorded concerning the number of participants for whom this information was available. Additionally, the interclass correlation (ICC) for the PCL total, factor and facet score was recorded as was the internal consistency (α) for the PCL total, factor and facet score as well as each risk assessment total score. Finally, the correlation (r) between the PCL total score, factor scores, and facet scores and each risk assessment measure was recorded.

Following the coding procedures, I scanned documents for potential sample overlap. When there were multiple publications or reports based on the same sample of data, I condensed these studies into a single sample and reported the data for each included risk assessment measure for the study with the largest sample size. If two sample sizes were the same, data was retained from the most recent article. Sample overlap was confirmed during correspondence with authors during the data request phase (see below). Using this procedure, 102 articles containing overlapping samples were

combined into 36 samples. Thus, altogether there were 211 samples eligible for inclusion following exclusion and combinations of overlapping samples.

Data Requests

For each eligible study, I contacted the corresponding author through electronic mail (email) using the contact information provided on each published study. Two authors were not contacted for additional information as all information of interest was included in the article and was obtained during the coding procedures. For dissertations and theses, contact information for the director or chair of the project was collected from the academic institution of interest and correspondence sent. First, each email included a brief summary of the purpose of the current study and listed the study or studies for which information was requested. Second, a table for reporting correlations between risk assessment measure and PCL total, factor, and facet scores was included. Any correlations reported in the article were already completed leaving blanks for correlations not reported. Finally, a list of additional information not presented in the study, including descriptive information for measures listed and demographic information for the sample was included. For multiple articles combined into a single study due to sample overlap a single email was sent requesting information for all risk measures reported across studies.

Once emails were sent, if an error or undeliverable message was received, up to date contact information was sought by conducting an internet search for more recent publications which may contain the individuals current contact information or a Google search for the author at academic institutions. For several authors, contact information was not available through academic or research institutions, however, authors had profiles on ResearchGate, a commercial social networking site for scientists and

researchers. In these cases, researchers were sent data requests through the direct messaging feature on this site. Following initial contact attempts, I allowed two months for researchers to respond to data requests. Twenty-six corresponding authors responded to my request with data for 29 samples. Two authors who responded to my request chose to send raw data sets or portions of raw data sets, on which I conducted the analyses requested. The final author response rate to my data request was 14%.

Following data collection, studies were retained for inclusion if data was available pertaining to at least one correlation between a PCL measure and a risk assessment measure. Only studies that reported no correlations between risk measures and the PCL measure and for which the authors did not provide additional information were excluded. Additionally, for studies including more than one risk assessment measure and information regarding their correlation with the PCL, a separate data entry was created. In total, 187 entries from 87 samples were included in the current study. Of those, 181 (96.8 %) contained a correlation between the PCL total score and risk measure total score. One hundred and eight (57.8%) contained a correlation between the PCL factor 1 score and risk measure total score. One hundred and nine (58.2%) contained a correlation between the PCL factor 2 score and risk measure total score. Fifty-nine (31.6%) , 57 (30.4%), 60 (32.1%) and 60 (32.1%) contained a correlation between the risk measure and PCL facet 1, facet 2, facet 3, and facet scores respectively (Table 1).

Table 1

Overview of Studies, Samples Sizes, and Correlations Included in the Current Review

Study/Sample	PCL Measure	Risk Measure	PCL Total Score		PCL Factor 1 Score		PCL Factor 2 Score		PCL Facet 1 Score		PCL Facet 2 Score		PCL Facet 3 Score		PCL Facet 4 Score	
			n	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*
Anderson et al, 2020	PCL-R	HCR-20	19	.94	19	.86	19	.88	19	.75	19	.90	19	.90	19	.65
Anderson et al, 2020	PCL:SV	HCR-20	14	.87												
Anderson et al, 2020	PCL-R	Static-99R	59	.56	59	.40	59	.62	59	.43	59	.31	59	.54	59	.63
Anderson et al, 2020	PCL:SV	Static-99R	33	.43												
Arbach-Lucioni et al, 2011	PCL:SV	HCR-20	78	.60	78	.50	78	.59								
Wakworth Sample	PCL-R	MASORR	409	.37	409	.36	409	.29								
Wakworth Sample	PCL-R	MnSOST-R	354	.30	354	.14	354	.35								
Wakworth Sample	PCL-R	RRASOR	442	.08	442	.03	442	.09								
Wakworth Sample	PCL-R	SORAG	442	.66	442	.28	442	.69								
Wakworth Sample	PCL-R	Static-2002	442	.36			442	.44								
Wakworth Sample	PCL-R	Static-99	442	.36	442	.13	442	.38								
Wakworth Sample	PCL-R	SVR-20	442	.58	442	.32	442	.55								

Study/Sample	PCL Measure	Risk Measure	PCL Total Score		PCL Factor 1 Score		PCL Factor 2 Score		PCL Facet 1 Score		PCL Facet 2 Score		PCL Facet 3 Score		PCL Facet 4 Score	
			n	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*
Wakworth Sample	PCL-R	VASOR	442	.35	442	.19	442	.34								
Wakworth Sample	PCL-R	VRAG	442	.70	442	.28	442	.74								
Beggs & Grace, 2008	PCL-R	Static-99	216	.40												
Brown et al, 2009	PCL-R	SIR-R1	136	-.57												
Caldwell et al, 2008	PCL:YV	J-SOAP-II	91	.07												
Caldwell et al, 2008	PCL:YV	JRAS	91	-.19												
Caldwell et al, 2008	PCL:YV	RRAS	91	-.18												
Caldwell et al, 2008	PCL:YV	TJSORAI	91	-.13												
Caldwell et al, 2008	PCL:YV	WDOC	91	-.06												
Caperton, 2005	PCL-R	MnSOST-R	1,983	.28												
Caperton, 2005	PCL-R	Static-99	1,983	.19												
Canadian Federal Sample	PCL-R	GSIR-R	106	.46	106	.08	106	.64								
Canadian Federal Sample	PCL-R	PRSF	106	.48	106	.16	106	.56								
Canadian Federal Sample	PCL-R	VSIR-R	106	.39	106	.19	106	.40								
Canadian Federal Sample	PCL-R	TTV	119	.66					119	.24	119	.14	119	.49	119	.86

Study/Sample	PCL Measure	Risk Measure	PCL Total Score		PCL Factor 1 Score		PCL Factor 2 Score		PCL Facet 1 Score		PCL Facet 2 Score		PCL Facet 3 Score		PCL Facet 4 Score	
			n	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*
Canadian Federal Sample	PCL-R	VRAG	119	.68	106	.21	106	.84								
Canadian Federal Sample	PCL-R	VRAG-R	119	.69	106	.18	106	.85								
Claix & Pham, 2004	PCL-R	HCR-20	86	.63	86	.51	86	.62								
PSIIS Sample	PCL-R	HCR-20	1,627	.80					1,709	.44	1,702	.49	1,696	.72	1,699	.70
PSIIS Sample	PCL-R	OGRS	1,632	.46					1,709	.05	1,702	.09	1,696	.49	1,699	.69
PSIIS Sample	PCL-R	RM2000/V	1,708	.51					1,709	.13	1,702	.11	1,696	.49	1,699	.71
PSIIS Sample	PCL-R	Static-99	354	.51					1,709	.32	1,702	.29	1,696	.39	1,699	.51
PSIIS Sample	PCL-R	SVR-20	340	.72					1,709	.51	1,702	.52	1,696	.57	1,699	.54
PSIIS Sample	PCL-R	VRAG	1,702	.72					1,709	.30	1,702	.29	1,696	.68	1,699	.79
Dahle, 2006	PCL-R	HCR-20	307	.76	307	.38	307	.79	307	.07	307	.47	307	.70	307	.68
Dahle, 2006	PCL-R	LSI-R	307	.61	307	.26	307	.69	307	-.04	307	.39	307	.68	307	.51
Barber-Rioja et al, 2012	PCL:SV	HCR-20	131	.72	131	.44	131		.75							
De Page et al, 2018	PCL-R	VRAG	72	.63	72	.29	72	.78								
De Page et al, 2018	PCL-R	SAPROF	72	-.55	72	-.46	72	-.53								
De Vogel et al, 2019	PCL-R	FAM	71	.59	71	.33	71	.58	71	.30	71	.29	71	.62	71	.39
De Vogel et al, 2019	PCL-R	HCR-20	71	.59	71	.32	71	.63	71	.17	71	.39	71	.62	71	.50
De Vogel et al, 2019	PCL-R	HCR-20V3	71	.60	71	.34	71	.64	71	.21	71	.40	71	.60	71	.54
De Vogel et al, 2019	PCL-R	START			67	.29	67	.42								

Study/Sample	PCL Measure	Risk Measure	PCL Total Score		PCL Factor 1 Score		PCL Factor 2 Score		PCL Facet 1 Score		PCL Facet 2 Score		PCL Facet 3 Score		PCL Facet 4 Score	
			n	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*
De Vogel et al, 2019	PCL-R	SAPROF	71	.47	71	-.38	71	-.41	70	-.20	71	-.47	71	-.41	71	-.31
Desmarais et al, 2012	PCL:SV	HCR-20	120	.73	118	.47	120	.71	119	.42	119	.39	120	.63	120	.57
Desmarais et al, 2012	PCL:SV	START/S	120	-.50	118	-.30	120	-.51	119	-.20	119	-.30	120	-.48	120	-.38
Desmarais et al, 2012	PCL:SV	START/V	120	.54	118	.34	120	.54	119	.29	119	.30	120	.55	120	.35
Dickson et al, 2013	PCL:SV	VRS	49	.52	49	.25	49	.56								
Dickson et al, 2013	PCL:SV	RoC*RoI	49	.11												
Dolan & Rennie, 2008	PCL:YV	SAVRY	99	.71					99	.42	99	.46	99	.65	99	.61
Nicholls et al, 2004 – Male sample	PCL:SV	VSC	146	.07	146	-.13	146	.01								
Nicholls et al, 2004 – Female sample	PCL:SV	VSC	90	.09	90	.09	90	.10								
Douglas et al, 2005	PCL:SV	HCR-20	556	.68	556	.44	556	.73								
Douglas et al, 2005b	PCL-R	HCR-20	188	.85	188	.56	188	.88								
Douglas et al, 2005b	PCL-R	VORAS	188	.41	188	.10	188	.51								
Douglas et al, 2005b	PCL-R	VRAG	188	.65	188	.32	188	.75								

Study/Sample	PCL Measure	Risk Measure	PCL Total Score		PCL Factor 1 Score		PCL Factor 2 Score		PCL Facet 1 Score		PCL Facet 2 Score		PCL Facet 3 Score		PCL Facet 4 Score	
			n	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*
Douglas et al, 2005b	PCL:SV	HCR-20	188	.81	188	.55	188	.86								
Douglas et al, 2005b	PCL:SV	VORAS	188	.36	188	.10	188	.53								
Douglas et al, 2005b	PCL:SV	VRAG	188	.60	188	.32	188	.72								
Douglass, 2009	PCL-R	HARM			39	-.36	39	-.34								
Douglass, 2009	PCL-R	HCR-20			39	.75	39	.78								
Douglass, 2009	PCL-R	VRAG			39	.78	39	.79								
Doyle et al, 2002	PCL:SV	VRAG	87	.81												
FECVSO Sample	PCL-R	DVRAG	65	.39												
FECVSO Sample	PCL-R	ODARA	65	.29												
FECVSO Sample	PCL-R	RRASOR	382	.24												
FECVSO Sample	PCL-R	SORAG	749	.78												
FECVSO Sample	PCL-R	Stable 2007	252	.49												
FECVSO Sample	PCL-R	Static-99	737	.53												
FECVSO Sample	PCL-R	Static-99R	181	.39												
FECVSO Sample	PCL-R	SVR-20	366	.77												
FECVSO Sample	PCL-R	VRAG	1,104	.77												
FECVSO Sample	PCL-R	VRS:SO	252	.50												
Farr, 2013	PCL:YV	J-SOAP-II	389	.80					389	.57	389	.52	389	.63	389	.70
Folino, 2015	PCL-R	HCR-20	153	.75	153	.45	153	.81	153	.21	153	.56	153	.73	153	.72
Folino, 2015	PCL-R	VRAG	153	.73	153	.39	153	.81	153	.20	153	.46	153	.66	153	.79
Fougere et al, 2015	PCL:SV	LS/CMI	72	.64												
Friesen, 1996	PCL:YV	YO-LSI	40	.55												
Grann & Wedin, 2002	PCL-R	SARA	88	.59	88	.38	88	.55								

Study/Sample	PCL Measure	Risk Measure	PCL Total Score		PCL Factor 1 Score		PCL Factor 2 Score		PCL Facet 1 Score		PCL Facet 2 Score		PCL Facet 3 Score		PCL Facet 4 Score	
			n	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*
Gray et al, 2003	PCL-R	HCR-20	34	.69	34	.49	34	.68								
Gray et al, 2004	PCL:SV	HCR-20	230	.78	221	.64	244	.74								
Gray et al, 2004	PCL:SV	OGRS	212	.35	212	.16	212	.51								
Haines et al, 2018 – Secure inpatient	PCL:SV	HCR-20	55	.71												
Haines et al, 2018 – Secure inpatient	PCL:SV	SAPROF	55	-.64												
Haines et al, 2018 – General inpatient	PCL:SV	HCR-20	100	.71												
Haines et al, 2018 – General inpatient	PCL:SV	SAPROF	100	-.66												
Haines et al, 2018 – Community	PCL:SV	HCR-20	106	.77												
Haines et al, 2018 – Community	PCL:SV	SAPROF	106	-.63												
Hausam Sample	PCL-R	LSI-R	274	.57	274	.19	274	.72	274	.10	274	.22	274	.62	274	.60
Hausam Sample	PCL-R	SAPROF	274	-.24	274	-.10	274	-.28	274	-.07	274	.10	274	-.300	274	-.17
Hill et al, 2012	PCL-R	HCR-20	139	.74												
Hill et al, 2012	PCL-R	Static-99	139	.41												
Hill et al, 2012	PCL-R	SVR-20	139	.75												
Hilterman et al, 2014	PCL:YV	SAVRY	105	.66												
Hilterman et al, 2014	PCL:YV	YLS/CMI	105	.74												
Hilton et al, 2008	PCL-R	DA	649	.36												
Hilton et al, 2008	PCL-R	DVRAG	649	.72												
Hilton et al, 2008	PCL-R	DVSI	649	.34												

Study/Sample	PCL Measure	Risk Measure	PCL Total Score		PCL Factor 1 Score		PCL Factor 2 Score		PCL Facet 1 Score		PCL Facet 2 Score		PCL Facet 3 Score		PCL Facet 4 Score	
			n	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*
Hilton et al, 2008	PCL-R	ODARA	649	.55												
Hilton et al, 2008	PCL-R	SARA	649	.55												
Hilton et al, 2008	PCL-R	VRAG	649	.72												
Hilton et al, 2016	PCL-R	HCR-20	63	.26												
Hilton et al, 2016	PCL-R	VRAG	63	.71												
Hogan & Olver, 2018	PCL-R	HCR-20V3	32	.81									32	.62	33	.81
Hogan & Olver, 2018	PCL-R	VRAG-R	32	.66									32	.58	33	.76
Hogan & Olver, 2018	PCL-R	VRS	32	.78									32	.77	33	.76
Hogan & Olver, 2016	PCL-R	HCR-20V3	77	.67					77	.22	77	.50	77	.72	72	.66
Holmqvist, 2008	PCL:SV	HCR-20	47	.63	47	.57	47	.40								
Jack, 2000	PCL:YV	YLS/CMI	149	.67	150	.45	149	.67								
Joyal et al, 2011	PCL-R	HCR-20	174	.84	173	.67	174	.80								
Kanters et al, 2017 – Child Sexual Abusers	PCL-R	SVR-20	27	.61					27	.52	27	.25	27	.58	27	.28
Kanters et al, 2017 – Rapists	PCL-R	SVR-20	35	.60					35	.33	35	.35	35	.49	35	.40
Kroner et al, 2007	PCL-R	GSIR							89	.31						
Kroner et al, 2007	PCL-R	LSI-R							89	.42						
Kroner et al, 2005	PCL-R	GSIR	206	.59												
Kroner et al, 2005	PCL-R	LSI-R	206	.77												

Study/Sample	PCL Measure	Risk Measure	PCL Total Score		PCL Factor 1 Score		PCL Factor 2 Score		PCL Facet 1 Score		PCL Facet 2 Score		PCL Facet 3 Score		PCL Facet 4 Score	
			n	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*
Kroner et al, 2005	PCL-R	VRAG	206	.73												
Kropp et al, 2011	PCL:SV	SAM	101	.46					101	.27	101	.50	101	.22	101	.32
Langton et al, 2009	PCL-R	HCR-20	42	.54	42	.50	42	.32	42	.36	42	.56	42	.11	42	.28
Langton et al, 2009	PCL-R	VRS	34	.42	34	.19	34	.46	34	.10	34	.25	34	.30	34	.41
Laxton, 1998	PCL-R	LSI-R	62	.83	62	.39	62	.89								
Laxton, 1998	PCL-R	SORAG	62	.90	62	.52	62	.87								
Lewis, 2004	PCL-R	VRS	123	.70	123	.37	123	.77								
Lewis & Ireland, 2019	PCL:SV	HCR-20	23	.65	11	.38	18	.70								
Lister, 2010	PCL-R	VRAG	94	.78	94	.58	94	.78	94	.51	94	.46	94	.67	94	.69
Loza & Simourd, 1994	PCL-R	LSI	161	.78	161	.53	161	.84								
Martinaki et al, 2013	PCL:SV	HCR-20	295	.61	295	.54	295	.57								
McCoy, 2015	PCL-R	LSI-R	241	.54	241	.38	241	.45	241	.21	241	.41	241	.36	241	.42
Khanna et al, 2014	PCL:YV	SAVRY	109	.72	109	.52	109	.72	109	.39	109	.51	109	.65	109	.63
Khanna et al, 2014	PCL:YV	YLS/CMI	109	.52	109	.37	109	.58	109	.33	109	.30	109	.56	109	.52
McDermott et al, 2011	PCL-R	COVR	146	.54												
McNiel et al, 2003	PCL:SV	HCR-20	100	.61	100	.51	100	.58								
McNiel et al, 2003	PCL:SV	VSC	100	.23	100	.27	100	.16								

Study/Sample	PCL Measure	Risk Measure	PCL Total Score		PCL Factor 1 Score		PCL Factor 2 Score		PCL Facet 1 Score		PCL Facet 2 Score		PCL Facet 3 Score		PCL Facet 4 Score	
			n	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*
Mills & Kroner, 2006	PCL-R	GSIR	209	.58												
Mills & Kroner, 2006	PCL-R	LSI-R	209	.76												
Mills & Kroner, 2006	PCL-R	VRAG	209	.74												
Morrissey et al, 2005	PCL-R	HCR-20	182	.54	182	.33	182	.65								
Morrissey et al, 2005	PCL-R	VRAG	202	.49	202	.28	202	.59								
Neal et al, 2015	PCL-R	HCR-20	230	.61	230	.42	230	.55	230	.46	230	.29	230	.45	230	.45
Neves et al, 2011	PCL-R	HCR-20	158	.75	158	.50	158	.75	158	.35	158	.50	158	.71	158	.61
Clearwater Sex Offender Sample	PCL-R	Static-99R	302	.38	302	.16	302	.43	302	.12	302	.16	302	.32	302	.44
Clearwater Sex Offender Sample	PCL-R	VRS:SO	302	.50	302	.39	302	.44	302	.32	302	.36	302	.40	302	.37
Nicholls, 2001 – Male Sample	PCL:SV	HCR-20	46	.72	46	.62	46	.64								
Nicholls, 2001 – Female Sample	PCL:SV	HCR-20	40	.71	40	.52	40	.69								
Nishinaka et al, 2016	PCL-R	HCR-20	71	.76	71	.47	71	.68	71	.38	71	.42	71	.67	71	.45
Oziel et al, 2020	PCL-R	SAPROF	50	-.21	50	-.14	50	-.22								
Parent et al, 2011	PCL-R	MnSOST-R	503	.51	503	.22	503	.54								
Parent et al, 2011	PCL-R	RM2000	503	.52	503	.14	503	.59								
Parent et al, 2011	PCL-R	RRASOR	503	.18	503	.02	503	.23								
Parent et al, 2011	PCL-R	SORAG	503	.77	503	.43	503	.77								

Study/Sample	PCL Measure	Risk Measure	PCL Total Score		PCL Factor 1 Score		PCL Factor 2 Score		PCL Facet 1 Score		PCL Facet 2 Score		PCL Facet 3 Score		PCL Facet 4 Score	
			n	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*
Parent et al, 2011	PCL-R	Static-2002	503	.41	503	.07	503	.50								
Parent et al, 2011	PCL-R	Static-99	503	.50	503	.19	503	.51								
Parent et al, 2011	PCL-R	SVR-20	503	.75	503	.58	503	.65								
Parent et al, 2011	PCL-R	VRAG	503	.76	503	.42	503	.76								
Penney et al, 2020	PCL-R	HCR-20	349	.54												
Simpson et al, 2015	PCL:SV	HCR-20	179	.85												
Pflueger et al, 2015	PCL:SV	HCR-20	258	.73	258	.59	258	.70	258	.42	258	.61	258	.67	258	.59
Pham et al, 2019	PCL-R	HCR-20	440	.43	424	.15	413	.52	381	.06	383	.14	377	.47	357	.48
Pham et al, 2019	PCL-R	VRAG	440	.44	424	.12	413	.52	381	.14	383	.04	377	.37	357	.60
Rennie, 2009	PCL:YV	SAVRY	135	.68					135	.38	135	.47	135	.58	135	.57
Rennie, 2009	PCL:YV	YLS/CMI	135	.43												
Schmidt Juvenile Sample	PCL:YV	SAVRY	133	.70	133	.34	133	.69	133	.07	133	.47	133	.62	133	.60
Schmidt Juvenile Sample	PCL:YV	YLS/CMI	133	.48	133	.23	133	.45	133	.09	133	.31	133	.41	133	.42
Rodriguez, Fernandez, & Gomez, 2015	PCL-R	VRAG	276	.55												
Sellbom et al, 2018	PCL:SV	HCR-20	99	.52	99	.27	99	.61	99	-.04	99	.45	99	.64	99	.31
Sellbom et al, 2018	PCL:SV	OaSys	99	.47	99	.19	99	.61	99	-.02	99	.31	99	.43	99	.53
Roth, 2005	PCL:YV	SAVRY	100	.74												
Shepherd et al, 2014	PCL:YV	SAVRY	213	.78	213	.49	213	.79	213	.39	213	.46	213	.69	213	.76

Study/Sample	PCL Measure	Risk Measure	PCL Total Score		PCL Factor 1 Score		PCL Factor 2 Score		PCL Facet 1 Score		PCL Facet 2 Score		PCL Facet 3 Score		PCL Facet 4 Score	
			n	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*
Shepherd et al, 2014	PCL:YV	YLS/CMI	213	.79	213	.48	213	.80	213	.38	213	.46	213	.71	213	.73
Sjostedt & Langstrom, 2002	PCL-R	RRASOR	51	.29	51	.25	51	.27	51	.23	51	.26	51	.20	51	.30
Sjostedt & Langstrom, 2002	PCL-R	SVR-20	51	.70	51	.53	51	.67	51	.53	51	.54	51	.65	51	.64
Sjostedt & Langstrom, 2002	PCL-R	VRAG	51	.63	51	.32	51	.73	51	.39	51	.32	51	.67	51	.69
St. Amand, 2002 – Sample 1	PCL-R	SIR-R1	157	.74	157	.47	157	.81								
St, Amand, 2002 – Sample 2	PCL-R	SIR-R1	233	.58	233	.15	233	.70								
Stockdale et al, 2014	PCL:YV	VRS-YV	147	.80					147	.39	147	.57	147	.70	147	.79
Sturup et al, 2016	PCL-R	HCR-20	91	.78												
Summers & Loza, 2004	PCL-R	VRAG	116	.80												
Thomson et al, 2008	PCL-R	VRAG	140	.73												
Tiegreen, 2009	PCL-R	VCRI	54	.38	54	.28	54	.48								
Viljoen et al, 2009	PCL:YV	Static-99	193	.35												
Viljoen et al, 2009	PCL:YV	ERASOR	193	.63												
Viljoen et al, 2009	PCL:YV	YLS/CMI	193	.77												
Vitacco et al, 2012	PCL:SV	VRAG	103	.35	103	.29	103	.31								
Wijetunga, 2015	PCL:YV	J-SOAP-II	143	.68	143	.44	143	.57	143	.03	143	.58	143	.49	143	.58

Study/Sample	PCL Measure	Risk Measure	PCL Total Score		PCL Factor 1 Score		PCL Factor 2 Score		PCL Facet 1 Score		PCL Facet 2 Score		PCL Facet 3 Score		PCL Facet 4 Score	
			n	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*
Wormith et al, 2007	PCL-R	LSI.CMI	61	.79	61	.58	61	.86	61	.42	61	.57	61	.65	61	.82
Zanatta, 2005	PCL-R	SORAG	164	.77	164	.30	164	.80								
Zanatta, 2005	PCL-R	Static-99	164	.72	164	.31	164	.70								
Zanatta, 2005	PCL-R	SVR-20	164	.39	164	-.02	164	.53								
Zanatta, 2005	PCL-R	VRAG	164	.78	164	.27	164	.83								
Zhu, Li, & Wang, 2016	PCL-R	HCR-20	75	.52	75	.28	75	.61	75	.21	75	.23	75	.57	75	.47
Texas SVP Sample	PCL-R	Static-99	746	.40	343	.19	344	.46	57	-.18	57	-.04	57	.26	61	.31
Texas SVP Sample	PCL-R	MnSOST-R	657	.45	355	.26	356	.51	51	.21	51	.10	51	.31	51	.55

Measures

PCL Family Measures

The PCL-R (Hare, 2003) is a 20-item checklist created to help in the assessment of the clinical construct of psychopathy, which is defined by a series of interpersonal, affective, and lifestyle characteristics developed for use with individuals aged 18 years and older. Each criterion is rated on a 3-point scale (0,1,2) depending on its presence in an individual, and these ratings are summative to obtain a total score within the range of 0-40. A cut-off of 30 is typically used to designate a diagnosis of psychopathy. Ratings for each item are based on a semi-structured interview, a review of collateral materials and interviews, and behavioral observations. In addition to the calculation of the total score, the PCL-R also contains two factors and four facets for which scores can be calculated. Factor 1 includes items on the interpersonal domain (facet 1), characterized by grandiosity, arrogance, callousness, superficiality, and manipulativeness, and affective domain (facet 2), characterized by short-temper, inability to form strong emotional bonds, and a lack of guilt. Factor 2 includes items on impulsive lifestyle (facet 3), including social deviance and impulsive behaviors, and antisocial behavior (facet 4), including violations of social convention (Hare, 1999). PCL-R administration takes between 150-180 minutes; 90-120 minutes of interview, and 60 minutes of collateral review.

The PCL:SV (Hart et al, 1995) was developed in response to the need for a shorter method of screening for psychopathy personality traits due to the length of time required to administer the PCL-R. The PCL:SV is a 12-item checklist created to help in the assessment of the clinical construct of psychopathy, adapted from the PCL-R. The PCL:SV is scored in the same manner as the PCL-R producing total scores within the

range of 0-24, and was developed for use with individuals aged 16 years and older. The PCL:SV can also be utilized to determine factor scores for the personality features of psychopathy (Factor 1) and behavioral patterns associated with psychopathy (Factor 2; Brazil & Forth, 2016). Unlike the PCL-R, the PCL:SV was designed to be administered within 50-90 minutes; 30-60 minutes of interview, and 20-30 minutes of collateral review.

The PCL-YV (Forth et al, 2003) was developed to aid in the assessment of psychopathic traits in adolescents and is developed for use on individuals ages 12-18 years. The PCL-YV is a 20-item checklist and contains the same factor and facet structure as the PCL-R. The PCL-YV produces scores ranging from 0-40 and is scored using the same scale as the PCL-R. PCL-YV administration takes between 150-180 minutes; 90-120 minutes of interview, and 60 minutes of collateral review.

Risk Assessment Measures

Together, studies selected for inclusion in this review reported correlations between a PCL measure and at least one of 53 different risk assessment measures (Table 2). Risk assessment measures selected for this review included both structured professional judgement measures and actuarial measures. Additionally, measures assessed for a variety of different types of risk including risk of violence, risk of sexual violence, risk of intimate partner violence, risk of general recidivism and risk of reconviction. The most commonly reported on studies included the HCR-20 (40 samples), VRAG (24 samples), Static-99 (10 samples), and SVR-20 (9 samples). One measure included in this review and subjected to separate analyses, the Structured

Assessment of Protective Factors (SAPROF), focused on the assessment of protective factors and was reported in 7 studies.

Table 2

Names, Types, and Risk Measured by the Risk Assessment Measures Included in the Study

Measure Name	Assessment Type	Type of Risk Measured	*K*
Historical Clinical Risk Management – 20, Version 2 (HCR-20)	SPJ	Violence	40
Violence Risk Appraisal Guide (VRAG)	Actuarial	Violence	24
Static-99	Actuarial	Sexual	10
Sexual Violence Risk – 20 (SVR-20)	SPJ	Sexual	9
Level of Service Inventory – Revised (LSI-R)	Actuarial	General	7
Structured Assessment of Protective Factors (SAPROF)	Protective Factors	Protective Factors	7
Structured Assessment of Violence Risk in Youth (SAVRY)	SPJ	Violence	7
Youth Level of Service/Case Management Inventory (YLS/CMI)	Actuarial	General	7
Sex Offender Risk Appraisal Guide (SORAG)	Actuarial	Sexual	5
Minnesota Sex Offender Screening Tool – Revied (MnSOST-R)	Actuarial	Sexual	4
Rapid Risk Assessment for Sexual Offense Recidvism (RRASOR)	Actuarial	Sexual	4
Static-99R	Actuarial	Sexual	4
Violence Risk Scale (VRS)	Actuarial	Violence	4
General Statistical Information of Recidivism Scale (GSIR)	Actuarial	General	3
Historical Clinical Risk Management – 20, Version 3 (HCR-20V3)	SPJ	Violence	3
Juvenile Sex Offender Assessment Protocol-II (J-SOAP-II)	Actuarial	Sexual	3
Statistical Informaiton on Recidivism – Revision 1 (SIR-R1)	Actuarial	General	3
Short-Term Assessment of Risk and Treatability (START)	SPJ	Violence	3
Violence Risk Scale: Sex Offender Version (VRS:SO)	Actuarial	Sexual	3
Violence Screening Checklist (VSC)	Actuarial	Inpatient violence risk	3

Measure Name	Assessment Type	Type of Risk Measured	*K*
Domestic Violence Risk Appraisal Guide (DVRAG)	Actuarial	Intimate partner violence	2
Ontario Domestic Assault Risk Assessment (ODARA)	Actuarial	Domestic violence	2
Offender Group Reconviction Scale (OGRS)	Actuarial	Reconviction	2
Risk Matrix 2000 (RM2000)	Actuarial	Violence and Sexual	2
Spousal Assault Risk Assessment (SARA)	SPJ	Violence	2
Static-2002	Actuarial	Sexual	2
Violent Offender Risk Assessment Scale (VORAS)	Actuarial	Violence	2
Violence Risk Appraisal Guide – Revised (VRAG-R)	Actuarial	Violence	2
Classification of Violence Risk (COVR)	SPJ	Violence	1
Danger Assessment (DA)	SPJ	Intimate partner violence	1
Domestic Violence Screening Instrument (DVSI)	Actuarial	Intimate partner violence	1
Estimate of Risk of Adolescent Sexual Offense Recidivism (ERASOR)	SPJ	Sexual	1
Female Additional Manual (FAM)	SPJ	Violence	1
General Statistical Information of Recidivism Scale – Revised (GSIR-R)	Actuarial	General	1
Hamilton Anatomy of Risk Management Tool (HARM)	SPJ	Violence	1
New Jersey Juvenile Risk Assessment Scale (JRAS)	Actuarial	Sexual	1
Level of Service Inventory (LSI)	Actuarial	General	1
Level of Service/Case Management Inventory	Actuarial	General	1
Multifactorial Assessment of Sex Offender Risk for Recidivism (MASORR)	SPJ	Sexual	1
Offender Assessment System (OASys)	Actuarial	Reconviction	1
Psychological Referral Screening Form (PRSF)	Unknown	Violence	1
New Jersey Registrant Risk Assessment Scale (RRAS)	Actuarial	Sexual	1
Risk of Reconviction x Risk of Reimprisonment (RoC*RoI)	Actuarial	Reconviction	1
Guidelines for Stalking Assessment and Management (SAM)	SPJ	Stalking	1
STABLE-2007	SPJ	Sexual	1

Measure Name	Assessment Type	Type of Risk Measured	*K*
Texas Juvenile Sex Offender Risk Assessment Instrument (TJSORAI)	Actuarial	Sexual	1
Two-Tiered Violence Risk Assessment (TTV)	Actuarial	Violence	1
Vermont Assessment of Sex Offender Risk (VASOR)	SPJ	Sexual	1
Violence Clinical Risk Indicator (VCRI)	Actuarial	Violence	1
Violence Risk Scale: Youth Version (VRS-YV)	Actuarial	Violence	1
Violent Statistical Information on Recidivism – Revised (VSIR-R)	Actuarial	Violence	1
Washington State Department of Corrections Measures (WDOC)	Unknown	Unknown	1
Youth Level of Service Inventory (YO-LSI)	Actuarial	General	1

Analytic Strategy

Measure of Effect Size

This meta-analysis used the correlation (*r*) value as the primary measure of effect size. Correlation values describe the strength of the relationship between two variables, that is to say, how much they vary in coordination with one another. Conventions for interpreting effect sizes of a correlation coefficient were developed by Cohen (1988). Cohen described a correlation coefficient of 0.10 as small, 0.30 as moderate, and 0.50 as large. For each study, correlation coefficients between risk assessment total scores and PCL total, factor, and facet scores were recorded. For studies utilizing structured professional judgment tools, correlation coefficients were calculated using pseudoactuarial coding methods (i.e., assigning a numerical value to each item and summing these items to develop a total score for the measure). Although peudoactuarial methods are not suggested for use with structured professional judgment measures in

clinical risk assessments, these calculations are common in risk assessment research to provide a method of direct comparison between actuarial and SPJ measures.

Prior to conducting data synthesis, correlation coefficients were converted to the Fisher's *z* scale, and variances calculated, using the Practical Meta-Analysis Effect Size Calculator available online as a companion to the Practical Meta-Analysis (Lipsey & Wilson, 2001). This conversion is conducted because correlation coefficients have limiting values (-1.00 and +1.00) and when estimating non-zero population values, may exhibit sampling distributions that are not normal. *Z*-scores, on the other hand have no limiting value and are normally distributed. As such, a conversation to *z*-scores should help to adjust these values to a normal distribution. Variances discerned during this procedure were then transformed into a standard error value in IBM SPSS Statistics Version 27 using the following formula (Borenstein et al, 2009);

$$SE_z = \sqrt{V_z}$$

The converted Fisher's *z* values were used for all analyses and transferred back to *r* values for reporting results.

Synthesis Methods

I used Jeffrey's Amazing Statistics Program (JASP) Version 0.16.3 (JASP Team, 2002) to conduct random-effects model meta-analyses using the Maximum Likelihood Model for the correlations between risk measures and the PCL total scores, factor scores, and facet scores as the effect sizes. Additional random effects meta-analyses were conducted separating data based on PCL version (i.e., PCL-R, PCL:YV, and PCL:SV), risk assessed (e.g., violence risk, sexual risk, domestic violence), and type of risk assessment measure (i.e., structured professional judgment, actuarial, or protective

factors). Finally, random-effect meta-analyses were conducted for each risk assessment measure for which at least 5 studies reported correlations with PCL scores.

Variation

The Cochran's Q and I^2 values were computed to aid in partitioning the variation within the meta-analysis. The Cochran's Q value is a weighted sum of square deviations of individual study effects from the pooled effect across studies. This functions as a measure of significance of heterogeneity across studies but does not give an estimate for the amount of variability between studies. The I^2 value, estimates the amount of the variance within the set of studies contained within the meta-analysis that is not due to sampling error, but due to studies or study features themselves.

Assessing for Publication Bias

Because of the decision to limit the literature search to published peer-reviewed research as well as master's theses and doctoral dissertations, consideration should be given to potential publication bias in the research reviewed. Null findings and low power findings are less likely to be published in journals in favor of studies with high powered statistical results. As such, the published research may result in skewed findings in a systematic review. A funnel plot was constructed to assess for publication bias. A funnel plot functions as a visual representation of the relationship between the study size and effect size. Studies with larger sample size will cluster towards the top of the plot and those with smaller sizes towards the bottom. Since smaller samples tend to have larger amounts of variance, they scatter more broadly giving the plot its funnel shape and name. In the absence of publication bias, studies are expected to be distributed symmetrically around the value of the mean effect size calculated during the meta-analysis. If there is

publication bias, scores will not be symmetrically distributed and will likely feature studies missing towards the middle and bottom of the funnel plot. The gaps in the funnel plot are likely to exist where nonsignificant results would be expected.

The trim-and-fill method was then employed to identify and correct for asymmetry which may arise in the funnel plot due to potential publication bias. This method first "trims" away smaller studies contributing to funnel plot asymmetry. It then reestablished the true center of the plot utilizing the data remaining before "filling" the plot by replacing studies that were removed as well as their potentially omitted counterparts (assuming for publication bias) around the newly established center. The trim and fill plot presents an estimate of the number of "missing" studies as well as an adjusted effect based on the inclusion of the filled studies. Two additional tests have been proposed to help quantify the amount of asymmetry observed in a funnel plot; the rank correlation test and the Egger's test. These values were calculated for each meta-analysis conducted.

Rosenthal's fail-safe N is a calculation that estimates how many studies with null effects would be needed and incorporated into an analyses before the results became non-significant. This calculation was devised to help ease concerns of researcher's that they would be missing studies with small or null effect sizes due to publication bias or the file-drawer effect (i.e., the tendency of papers will small or null effects to be rejected for publication in favor of studies with larger effect sizes or more impressive results). If the fail-safe N suggests that only a small number of missing studies would render the results non-significant, then researchers should be concerned about the effects of publication bias on the results of their analyses. However, if a large amount of studies would be

required to nullify results, this calculation suggests that there should be less concern about missing data.

Moderator Analyses

In addition to the meta-analyses mentioned above, meta-regression using the same methodology were run with moderator variables including the type of population utilized (i.e., correctional, institutional, outpatient, etc.), demographic information (sex, age, and race/ethnicity), the country in which the study was conducted, the year the study was published or completed (for theses and dissertations), the type of report from which the sample came (i.e., journals, theses, and dissertations), method of scoring of the PCL measure (i.e., File and interview or file only), and reliability statistics (alpha). Continuous moderation was conducted for age of participants (i.e., mean age of the sample), race/ethnicity (i.e., % of the sample Caucasian), year of publication and reliability coefficients. Categorical moderation was conducted for age of participants (i.e., adult v. juvenile samples), population utilized, sex of the sample (i.e., male v. female samples), country of the study, type of publication, and PCL scoring method. For combined samples, the earliest publication date was selected. Additionally, if one of the studies in the sample was published and others were theses/dissertations, academic journal was selected as the source for the data. For these collapsed data, the largest sample size, or the sample size utilized for the correlations, was used to determine demographic information for the sample. These analyses were conducted to help determine if the moderators mentioned above may be contributing to the effect sizes observed in the meta-analyses conducted.

CHAPTER III

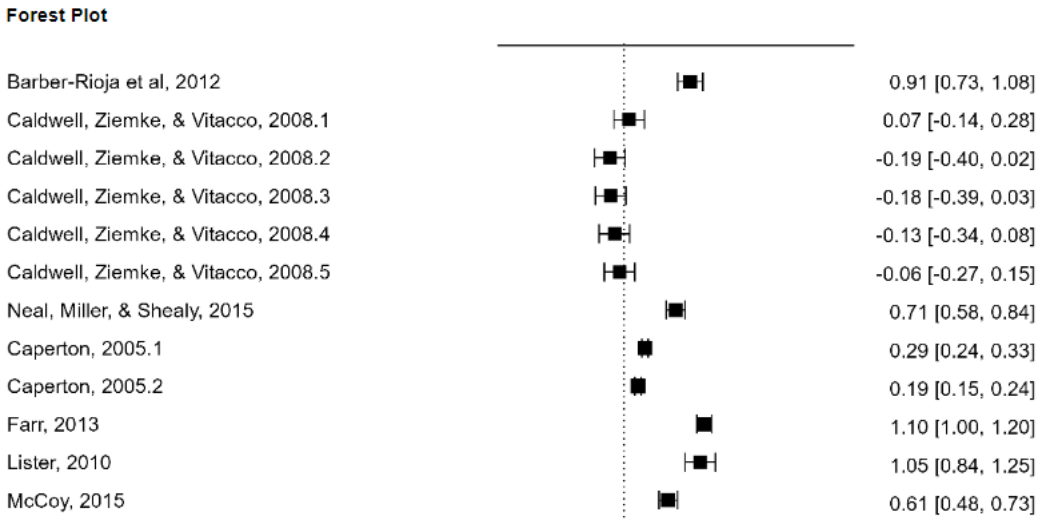
Results

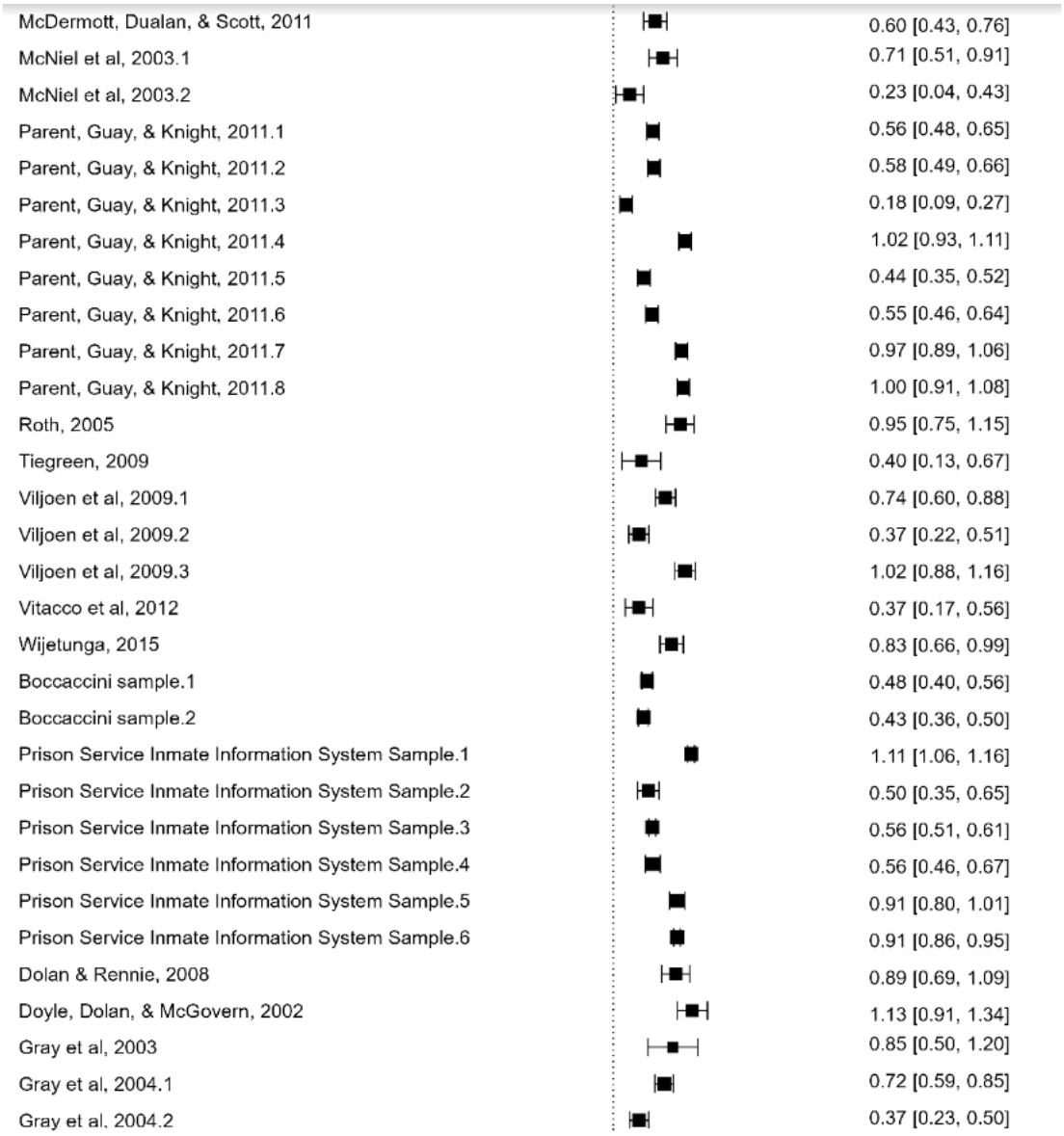
Comparison of PCL Total, Factor, and Facet Scores

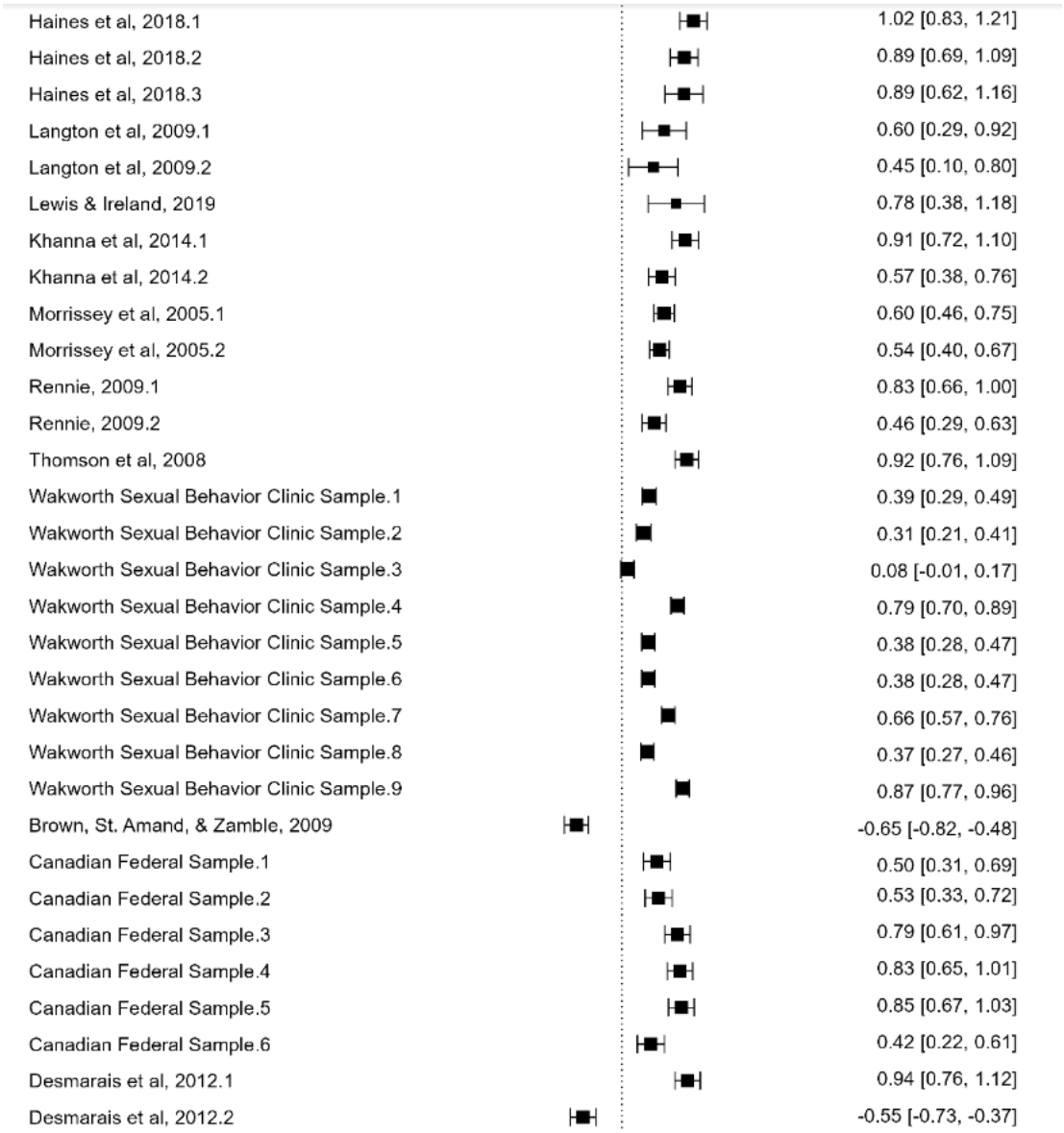
My first question was whether there was any evidence of overlap between risk measure total scores and PCL total scores. One hundred and eighty-one correlations between PCL total score and risk assessment measure total scores were analyzed (Figure 2). The mean effect size for PCL-R total score and risk assessment measure total scores was r=.59, which was moderate to large in size and statistically significant (k=181, 95% CI = .56-.62; Z = 26.79, p < .001; see Table 3). Variability statistics suggest there is more variability among the included study effects than is attributable to sampling error alone (Q(180) = 4,638.80, p<.001; I^2 = 96.39). This variability is likely better accounted for by between sample differences supporting the need for additional moderator analyses attempting to identify sources of these differences.

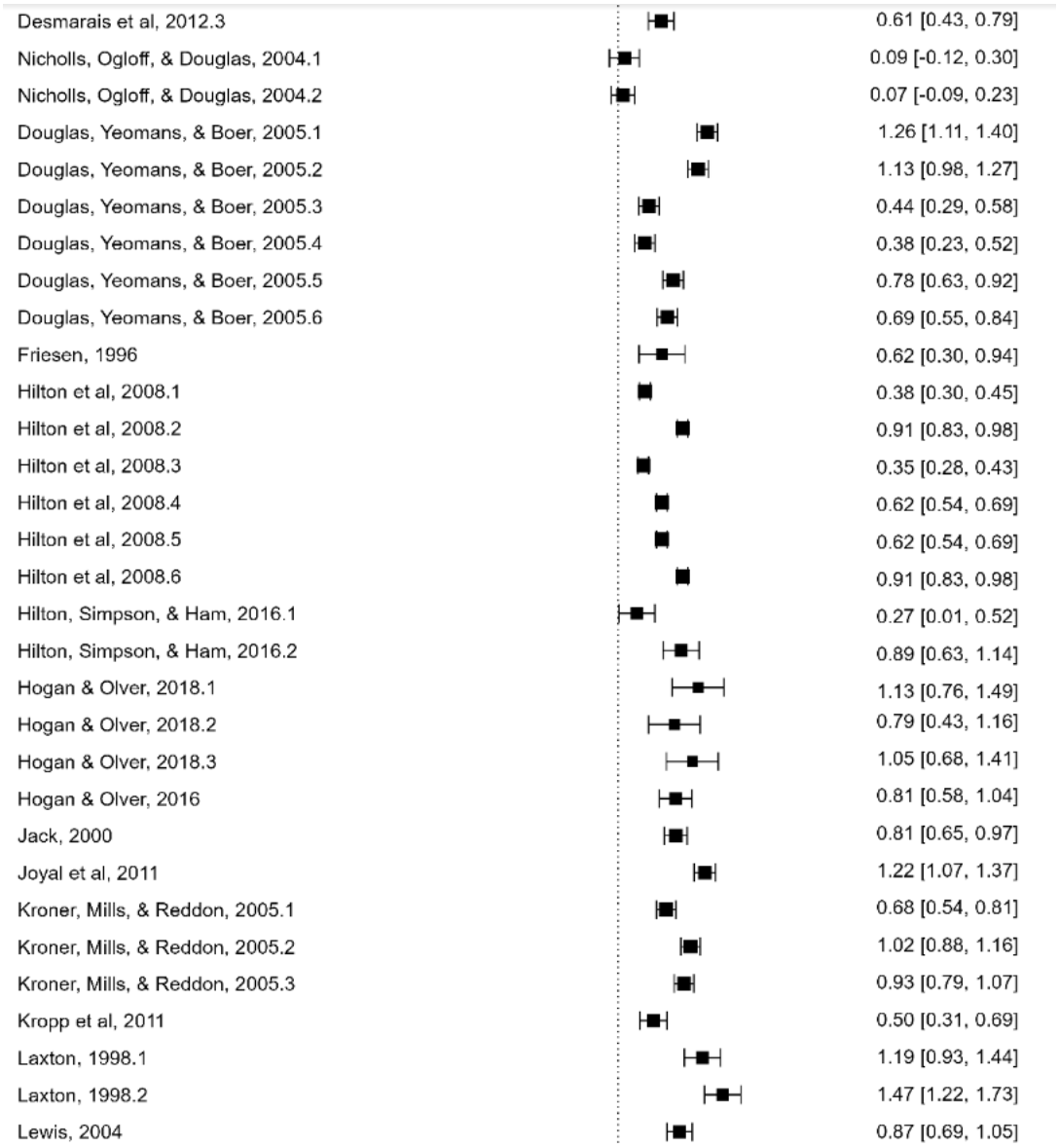
Figure 2

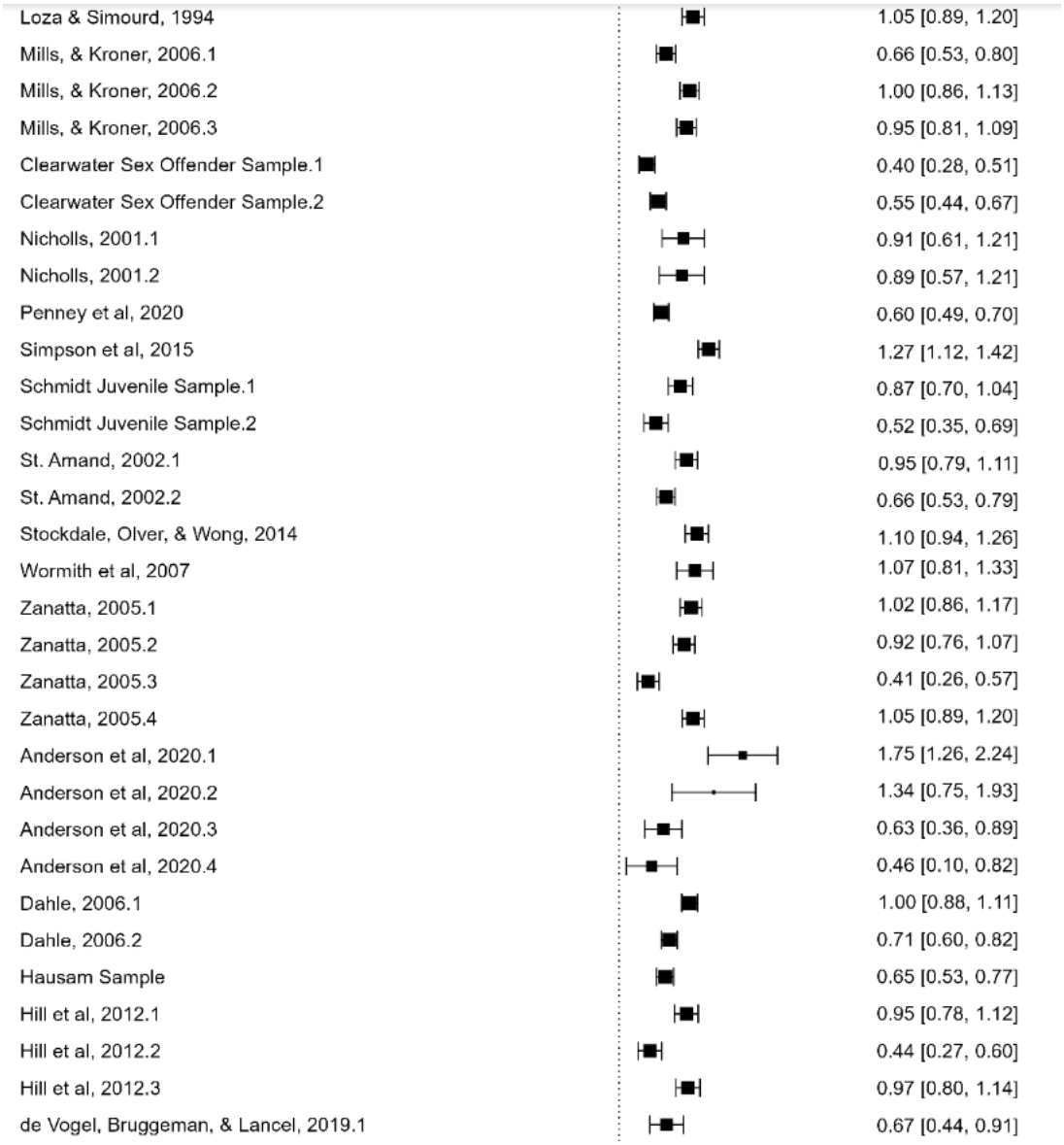
Forest Plot of Studies Reporting PCL Total Score Correlation with Risk Assessment Measures

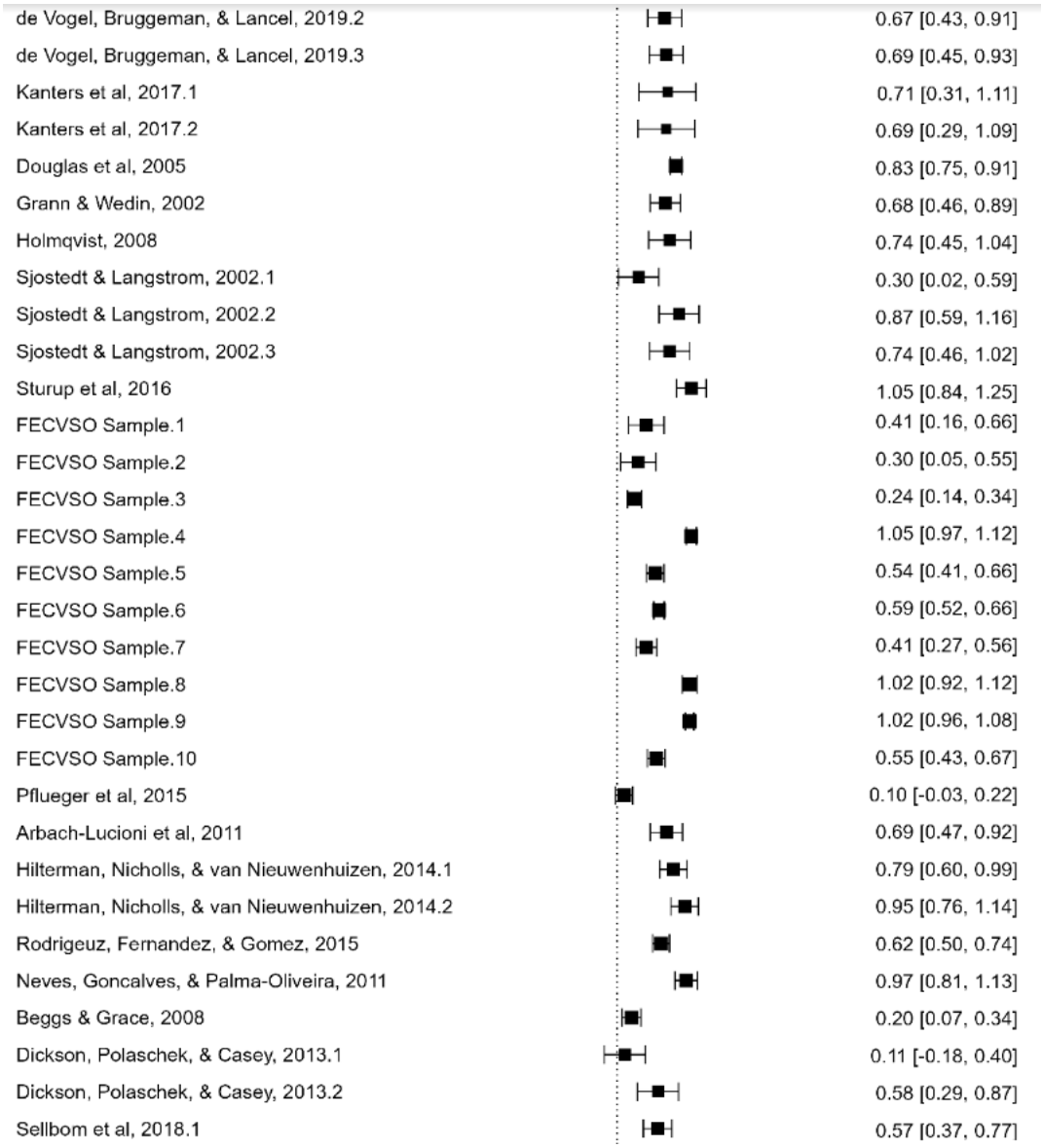



McDermott, Dualan, & Scott, 2011 0.60 [0.43, 0.76]
McNiel et al, 2003.1 0.71 [0.51, 0.91]
McNiel et al, 2003.2 0.23 [0.04, 0.43]
Parent, Guay, & Knight, 2011.1 0.56 [0.48, 0.65]
Parent, Guay, & Knight, 2011.2 0.58 [0.49, 0.66]
Parent, Guay, & Knight, 2011.3 0.18 [0.09, 0.27]
Parent, Guay, & Knight, 2011.4 1.02 [0.93, 1.11]
Parent, Guay, & Knight, 2011.5 0.44 [0.35, 0.52]
Parent, Guay, & Knight, 2011.6 0.55 [0.46, 0.64]
Parent, Guay, & Knight, 2011.7 0.97 [0.89, 1.06]
Parent, Guay, & Knight, 2011.8 1.00 [0.91, 1.08]
Roth, 2005 0.95 [0.75, 1.15]
Tiegreen, 2009 0.40 [0.13, 0.67]
Viljoen et al, 2009.1 0.74 [0.60, 0.88]
Viljoen et al, 2009.2 0.37 [0.22, 0.51]
Viljoen et al, 2009.3 1.02 [0.88, 1.16]
Vitacco et al, 2012 0.37 [0.17, 0.56]
Wijetunga, 2015 0.83 [0.66, 0.99]
Boccaccini sample.1 0.48 [0.40, 0.56]
Boccaccini sample.2 0.43 [0.36, 0.50]
Prison Service Inmate Information System Sample.1 1.11 [1.06, 1.16]
Prison Service Inmate Information System Sample.2 0.50 [0.35, 0.65]
Prison Service Inmate Information System Sample.3 0.56 [0.51, 0.61]
Prison Service Inmate Information System Sample.4 0.56 [0.46, 0.67]
Prison Service Inmate Information System Sample.5 0.91 [0.80, 1.01]
Prison Service Inmate Information System Sample.6 0.91 [0.86, 0.95]
Dolan & Rennie, 2008 0.89 [0.69, 1.09]
Doyle, Dolan, & McGovern, 2002 1.13 [0.91, 1.34]
Gray et al, 2003 0.85 [0.50, 1.20]
Gray et al, 2004.1 0.72 [0.59, 0.85]
Gray et al, 2004.2 0.37 [0.23, 0.50]


Haines et al, 2018.1 1.02 [0.83, 1.21]
Haines et al, 2018.2 0.89 [0.69, 1.09]
Haines et al, 2018.3 0.89 [0.62, 1.16]
Langton et al, 2009.1 0.60 [0.29, 0.92]
Langton et al, 2009.2 0.45 [0.10, 0.80]
Lewis & Ireland, 2019 0.78 [0.38, 1.18]
Khanna et al, 2014.1 0.91 [0.72, 1.10]
Khanna et al, 2014.2 0.57 [0.38, 0.76]
Morrissey et al, 2005.1 0.60 [0.46, 0.75]
Morrissey et al, 2005.2 0.54 [0.40, 0.67]
Rennie, 2009.1 0.83 [0.66, 1.00]
Rennie, 2009.2 0.46 [0.29, 0.63]
Thomson et al, 2008 0.92 [0.76, 1.09]
Wakworth Sexual Behavior Clinic Sample.1 0.39 [0.29, 0.49]
Wakworth Sexual Behavior Clinic Sample.2 0.31 [0.21, 0.41]
Wakworth Sexual Behavior Clinic Sample.3 0.08 [-0.01, 0.17]
Wakworth Sexual Behavior Clinic Sample.4 0.79 [0.70, 0.89]
Wakworth Sexual Behavior Clinic Sample.5 0.38 [0.28, 0.47]
Wakworth Sexual Behavior Clinic Sample.6 0.38 [0.28, 0.47]
Wakworth Sexual Behavior Clinic Sample.7 0.66 [0.57, 0.76]
Wakworth Sexual Behavior Clinic Sample.8 0.37 [0.27, 0.46]
Wakworth Sexual Behavior Clinic Sample.9 0.87 [0.77, 0.96]
Brown, St. Amand, & Zamble, 2009 -0.65 [-0.82, -0.48]
Canadian Federal Sample.1 0.50 [0.31, 0.69]
Canadian Federal Sample.2 0.53 [0.33, 0.72]
Canadian Federal Sample.3 0.79 [0.61, 0.97]
Canadian Federal Sample.4 0.83 [0.65, 1.01]
Canadian Federal Sample.5 0.85 [0.67, 1.03]
Canadian Federal Sample.6 0.42 [0.22, 0.61]
Desmarais et al, 2012.1 0.94 [0.76, 1.12]
Desmarais et al, 2012.2 -0.55 [-0.73, -0.37]


Desmarais et al, 2012.3 0.61 [0.43, 0.79]
Nicholls, Ogloff, & Douglas, 2004.1 0.09 [-0.12, 0.30]
Nicholls, Ogloff, & Douglas, 2004.2 0.07 [-0.09, 0.23]
Douglas, Yeomans, & Boer, 2005.1 1.26 [1.11, 1.40]
Douglas, Yeomans, & Boer, 2005.2 1.13 [0.98, 1.27]
Douglas, Yeomans, & Boer, 2005.3 0.44 [0.29, 0.58]
Douglas, Yeomans, & Boer, 2005.4 0.38 [0.23, 0.52]
Douglas, Yeomans, & Boer, 2005.5 0.78 [0.63, 0.92]
Douglas, Yeomans, & Boer, 2005.6 0.69 [0.55, 0.84]
Friesen, 1996 0.62 [0.30, 0.94]
Hilton et al, 2008.1 0.38 [0.30, 0.45]
Hilton et al, 2008.2 0.91 [0.83, 0.98]
Hilton et al, 2008.3 0.35 [0.28, 0.43]
Hilton et al, 2008.4 0.62 [0.54, 0.69]
Hilton et al, 2008.5 0.62 [0.54, 0.69]
Hilton et al, 2008.6 0.91 [0.83, 0.98]
Hilton, Simpson, & Ham, 2016.1 0.27 [0.01, 0.52]
Hilton, Simpson, & Ham, 2016.2 0.89 [0.63, 1.14]
Hogan & Olver, 2018.1 1.13 [0.76, 1.49]
Hogan & Olver, 2018.2 0.79 [0.43, 1.16]
Hogan & Olver, 2018.3 1.05 [0.68, 1.41]
Hogan & Olver, 2016 0.81 [0.58, 1.04]
Jack, 2000 0.81 [0.65, 0.97]
Joyal et al, 2011 1.22 [1.07, 1.37]
Kroner, Mills, & Reddon, 2005.1 0.68 [0.54, 0.81]
Kroner, Mills, & Reddon, 2005.2 1.02 [0.88, 1.16]
Kroner, Mills, & Reddon, 2005.3 0.93 [0.79, 1.07]
Kropp et al, 2011 0.50 [0.31, 0.69]
Laxton, 1998.1 1.19 [0.93, 1.44]
Laxton, 1998.2 1.47 [1.22, 1.73]
Lewis, 2004 0.87 [0.69, 1.05]


Loza & Simourd, 1994 1.05 [0.89, 1.20]
Mills, & Kroner, 2006.1 0.66 [0.53, 0.80]
Mills, & Kroner, 2006.2 1.00 [0.86, 1.13]
Mills, & Kroner, 2006.3 0.95 [0.81, 1.09]
Clearwater Sex Offender Sample.1 0.40 [0.28, 0.51]
Clearwater Sex Offender Sample.2 0.55 [0.44, 0.67]
Nicholls, 2001.1 0.91 [0.61, 1.21]
Nicholls, 2001.2 0.89 [0.57, 1.21]
Penney et al, 2020 0.60 [0.49, 0.70]
Simpson et al, 2015 1.27 [1.12, 1.42]
Schmidt Juvenile Sample.1 0.87 [0.70, 1.04]
Schmidt Juvenile Sample.2 0.52 [0.35, 0.69]
St. Amand, 2002.1 0.95 [0.79, 1.11]
St. Amand, 2002.2 0.66 [0.53, 0.79]
Stockdale, Olver, & Wong, 2014 1.10 [0.94, 1.26]
Wormith et al, 2007 1.07 [0.81, 1.33]
Zanatta, 2005.1 1.02 [0.86, 1.17]
Zanatta, 2005.2 0.92 [0.76, 1.07]
Zanatta, 2005.3 0.41 [0.26, 0.57]
Zanatta, 2005.4 1.05 [0.89, 1.20]
Anderson et al, 2020.1 1.75 [1.26, 2.24]
Anderson et al, 2020.2 1.34 [0.75, 1.93]
Anderson et al, 2020.3 0.63 [0.36, 0.89]
Anderson et al, 2020.4 0.46 [0.10, 0.82]
Dahle, 2006.1 1.00 [0.88, 1.11]
Dahle, 2006.2 0.71 [0.60, 0.82]
Hausam Sample 0.65 [0.53, 0.77]
Hill et al, 2012.1 0.95 [0.78, 1.12]
Hill et al, 2012.2 0.44 [0.27, 0.60]
Hill et al, 2012.3 0.97 [0.80, 1.14]
de Vogel, Bruggeman, & Lancel, 2019.1 0.67 [0.44, 0.91]


de Vogel, Bruggeman, & Lancel, 2019.2
0.67 [0.43, 0.91]
de Vogel, Bruggeman, & Lancel, 2019.3
0.69 [0.45, 0.93]
Kanters et al, 2017.1
0.71 [0.31, 1.11]
Kanters et al, 2017.2
0.69 [0.29, 1.09]
Douglas et al, 2005
0.83 [0.75, 0.91]
Grann & Wedin, 2002
0.68 [0.46, 0.89]
Holmqvist, 2008
0.74 [0.45, 1.04]
Sjostedt & Langstrom, 2002.1
0.30 [0.02, 0.59]
Sjostedt & Langstrom, 2002.2
0.87 [0.59, 1.16]
Sjostedt & Langstrom, 2002.3
0.74 [0.46, 1.02]
Sturup et al, 2016
1.05 [0.84, 1.25]
FECVSO Sample.1
0.41 [0.16, 0.66]
FECVSO Sample.2
0.30 [0.05, 0.55]
FECVSO Sample.3
0.24 [0.14, 0.34]
FECVSO Sample.4
1.05 [0.97, 1.12]
FECVSO Sample.5
0.54 [0.41, 0.66]
FECVSO Sample.6
0.59 [0.52, 0.66]
FECVSO Sample.7
0.41 [0.27, 0.56]
FECVSO Sample.8
1.02 [0.92, 1.12]
FECVSO Sample.9
1.02 [0.96, 1.08]
FECVSO Sample.10
0.55 [0.43, 0.67]
Pflueger et al, 2015
0.10 [-0.03, 0.22]
Arbach-Lucioni et al, 2011
0.69 [0.47, 0.92]
Hilterman, Nicholls, & van Nieuwenhuizen, 2014.1
0.79 [0.60, 0.99]
Hilterman, Nicholls, & van Nieuwenhuizen, 2014.2
0.95 [0.76, 1.14]
Rodrigeuz, Fernandez, & Gomez, 2015
0.62 [0.50, 0.74]
Neves, Goncalves, & Palma-Oliveira, 2011
0.97 [0.81, 1.13]
Beggs & Grace, 2008
0.20 [0.07, 0.34]
Dickson, Polaschek, & Casey, 2013.1
0.11 [-0.18, 0.40]
Dickson, Polaschek, & Casey, 2013.2
0.58 [0.29, 0.87]
Sellbom et al, 2018.1
0.57 [0.37, 0.77]

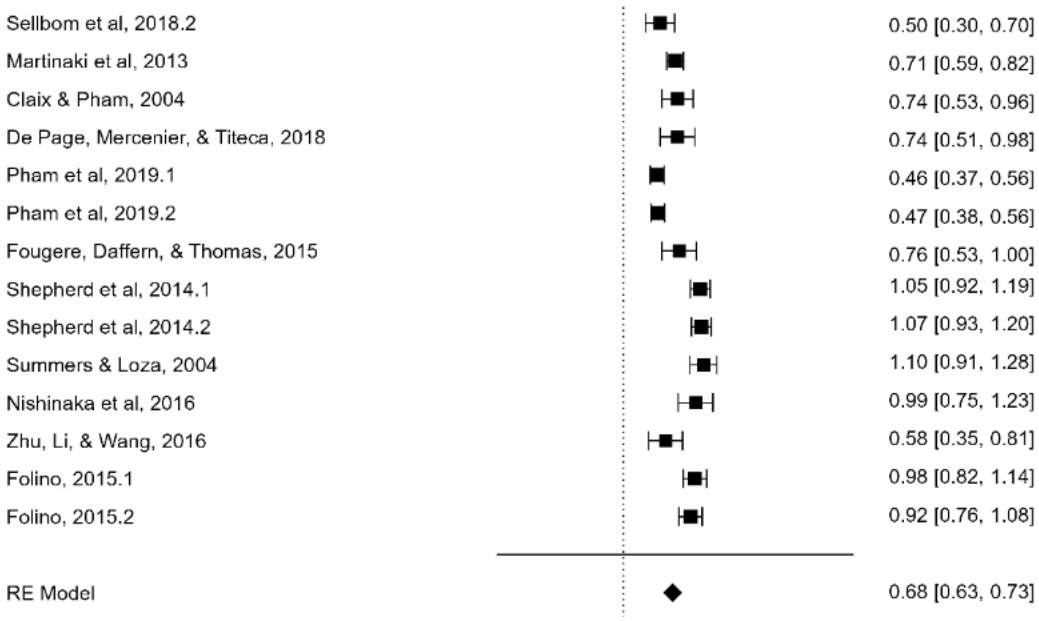


Table 3

Overall Meta-Analytic Effects and Measures of Heterogeneity for PCL Scores

PCL Score	r	95% CI	SE	*k*	Total N	*Q*	I^2 (%)
Total	.59***	.56-.62	.03	181	47,779	4,632.80***	96.39
Factor 1	.34***	.30-.38	.02	107	20,658	834.77***	88.09
Factor 2	.63***	.59-.66	.03	108	21,116	1,949.06***	95.21
Facet 1	.27***	.22-.32	.03	58	17,545	490.73***	87.63
Facet 2	.37***	.32-.41	.03	56	17,329	585.79***	88.98
Facet 3	.56***	.55-.60	.03	59	17,380	663.94***	93.06
Facet 4	.59***	.53-.62	.04	59	17,360	776.02***	93.72

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

Mean effect sizes were calculated to assess for evidence of overlap between risk measure total scores and PCL Factor scores. The mean effect size for PCL-R Factor scores and risk assessment measure total scores was r=.34 for Factor 1 and r =.63 for Factor 2, which were statistically significant (k = 107, 95% CI = .30-.38; Z = 16.56, p < .001; k = 108, 95% CI = .59-.66; Z = 22.64, p < .001) respectively. Similar to the analyses for correlations with PCL total scores, there was considerable variability among study effect sizes between PCL Factor scores and risk assessment total scores, Q(106) =

834.77, $p<.001$, $I^2 = 88.09$ for Factor 1 correlations and $Q(107) = 1{,}949.06$, $p<.001$, $I^2 = 95.21$ for Factor 2 correlations.

Finally, mean effect sizes were also calculated to assess for overlap between risk measure total scores and PCL facet scores. The mean effect size for PCL-R facet 1 score and risk assessment measure total scores was $r=.27$, which was statistically significant ($k = 58$, 95% CI = .22-.32; Z = 10.63, $p < .001$). For the association between the PCL-R facet 2 score and risk assessment measure total scores the mean effect size was $r = .37$, which was statistically significant ($k = 56$, 95% CI = .32-.41, Z = 13.96, $p < .001$). For the association with facet 3, the mean effect size was $r = .56$ and for facet 4 the mean effect size was $r = .59$. Both associations were statistically significant ($k = 59$, 95% CI = .55-.60; Z = 18.49, $p<.001$ and $k = 59$, 95% CI = .53-.62; Z = 18.46, $p < .001$ respectively). Considerable amount of heterogeneity not explained by sampling error was indicated for correlations between PCL total scores and all four facet scores ($Q(57) = 490.73$, $p<.001$, $I^2 = 87.63$; $Q(55) = 585.79$, $p<.001$, $I^2 = 88.98$; $Q(58) = 663.94$, $p<.001$, $I^2 = 93.06$; and $Q(58) = 776.02$, $p<.001$, $I^2 = 93.72$ for correlations with facet 1, 2, 3, and 4 respectively).

Comparison of Studies Based on PCL Measure Used

The next series of meta-analyses focused on exploring whether the mean effect sizes for correlations between PCL measure and risk assessment measures depends on the version of the PCL (e.g., PCL-R, PCL:SV, PCL:YV). To explore this question, three separate meta-analyses were run, one featuring all of the studies reporting risk assessment correlations with the PCL-R, one featuring all studies reporting risk assessment correlations with the PCL:SV, and one featuring studies reporting risk assessment

correlations with the PCL:YV. Overall effects and measures of heterogeneity for correlations between risk measures and the total, factor, and facet scores for each version of the PCL are noted in Table 4.

Mean effects for PCL total, factor, and facet scores were statistically significant regardless of PCL measure utilized. Overall, the effect size for the correlation between PCL total scores and risk measure scores was in the large range (*r*s = .56-.60), while the effect size for the correlation with Factor 1 scores was in the moderate range (*r*s = .32-.42)for all PCL measures. Although the Factor 2 effect sizes fell within the large range (*r*s = .55-.68) for all PCL measures, the magnitude of association between PCL:SV Factor 2 scores and overall risk was smaller than the effects for PCL-R and PCL:YV scores. Similarly, PCL:SV effect sizes were smaller for facet 1 (small range), facet 3 (moderate range), and facet 4 (moderate range) scores than effect sizes for correlations with the PCL-R and PCL:YV (moderate range for facet 1, and large range for facets 3 and 4 effect sizes).

Variability for correlations between risk measure total scores and the PCL:YV Factor 1 and facet 2 score suggest only small amounts of variability among the included study effects which are most likely attributable to sampling error. However, for all other correlations with PCL:YV total, factor, and facet scores, as well as all analyses for correlation with the PCL-R and PCL:SV scores, a significant amount of variability among the included study effects was noted. In these cases, variability was likely not solely attributable to sampling error, but also includes significant between study variability.

Table 4

Overall Meta-Analytic Effects and Measures of Heterogeneity, Correlations Based on PCL Version

	r	95% CI	SE	*k*	Total N	*Q*	I^2 (%)
PCL-R							
Total Score	.60***	.56-.63	.03	122	39,866	5,817.17***	99.75
Factor 1 Score	.32***	.28-.36	.02	75	15,856	941.13***	98.48
Factor 2 Score	.64***	.59-.68	.04	76	16,279	2,177.69***	99.47
Facet 1 Score	.27***	.21-.32	.03	41	14,808	766.61***	91.56
Facet 2 Score	.33***	.28-.39	.03	39	14,592	698.16***	92.32
Facet 3 Score	.56***	.51-.60	.04	42	14,640	936.24***	94.18
Facet 4 Score	.59***	.53-.63	.04	42	14,620	1,145.11***	95.14
PCL:SV							
Total Score	.56***	.46-.64	.06	35	4,422	598.65***	94.12
Factor 1 Score	.36***	.26-.45	.05	25	3,599	231.86***	89.39
Factor 2 Score	.55***	.44-.65	.08	25	3,635	480.28***	94.96
Facet 1 Score	.18*	.00-.34	.09	7	914	52.36***	86.01
Facet 2 Score	.34**	.13-.52	.12	7	914	88.44***	91.29
Facet 3 Score	.42**	.12-.64	.16	7	917	168.31***	95.79
Facet 4 Score	.35**	.11-.55	.13	7	917	106.39***	93.18
PCL:YV							
Total Score	.56***	.43-.67	.09	24	3,491	502.94***	95.79
Factor 1 Score	.42***	.36-.48	.04	8	1,203	12.01	32.25
Factor 2 Score	.68***	.59-.75	.07	8	1,202	50.32***	83.01
Facet 1 Score	.33***	.22-.42	.06	11	1,823	69.62***	91.56
Facet 2 Score	.47***	.43-.52	.03	11	1,823	16.79	31.63
Facet 3 Score	.62***	.57-.66	.04	11	1,823	28.26**	62.55
Facet 4 Score	.64***	.58-.70	.05	11	1,823	53.55***	80.24

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

Comparison of Studies Based on Type of Risk Assessment

Next, I wanted to explore whether there was any evidence of overlap between risk measure total scores and PCL scores for risk measure based on the type of risk measure (i.e., actuarial or structured professional judgment). First, I investigated potential overlap between actuarial risk measure total scores and PCL scores. The mean effect for PCL total score and actuarial risk measure total score was r=.55 which was statistically significant (k=108, 95% CI = .50-.59; Z = 18.02, p < .001). Variability statistics suggest there is more variability among the included study effects than is attributable to sampling

error alone ($Q(102) = 2{,}994.51$, $p<.001$; $I^2 = 97.02$). Similarly, mean effect sizes for actuarial risk total scores was larger for Factor 2 than Factor 1 (Table 5). At the facet level, both facets 3 and 4 demonstrated large effect sizes, and the smallest effect size was observed for facet 1.

Variability statistics suggest there is more variability among the included study effects than is attributable to sampling error alone for PCL total, factor, and facet score calculations (Table 5). This variability is likely better accounted for by between sample differences supporting the need for additional analyses aimed at identifying some of the sources of these differences.

Table 5

Overall Meta-Analytic Effects and Measures of Heterogeneity, Actuarial Risk Measures

	r	95% CI	SE	*k*	Total N	*Q*	I^2 (%)
Correlation, Total Score	.55***	.50-.59	.03	108	34,037	2,994.51***	97.02
Correlation, Factor 1	.28***	.24-.32	.02	59	13,074	328.23***	82.73
Correlation, Factor 2	.63***	.57-.68	.04	60	13,506	1,189.77***	95.69
Correlation, Facet 1	.25***	.19-31	.03	28	10,784	237.08***	88.52
Correlation, Facet 2	.33***	.27-.39	.04	26	10,580	259.17***	89.25
Correlation, Facet 3	.54***	.48-.59	.04	28	10,614	262.82***	90.21
Correlation, Facet 4	.63***	.57-.68	.05	28	10,612	446.99***	94.72

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

Second, I investigated potential overlap between structured professional judgment risk measure scores and PCL scores (Table 6). The mean effect for PCL total score and SPJ risk measure total score was r=.65, somewhat larger than for actuarial risk measures, which was statistically significant (k= 70, 95% CI = .61-.69; Z = 21.44, $p < .001$). Variability statistics suggest there is more variability among the included study effects than is attributable to sampling error alone ($Q(69) = 1{,}210.47$, $p<.001$; $I^2 = 93.78$). This variability likely better accounted for by between sample differences supporting the need

for additional analyses considering some of these differences. Analyses between SPJ total scores and PCL factor and facet scores also yielded statistically significant mean effects and high amounts of variability (Table 6). A similar pattern of associations emerged. Effect sizes were large for PCL-R Factor 2 and facets 3 and 4. Interestingly, the effect sizes for the association between SJP measures and PCL-R Factor 1 and facet 2 were moderate; however, the effect size for facet 1 remained modest.

Table 6

Overall Meta-Analytic Effects and Measures of Heterogeneity, SPJ Risk Measures

	r	95% CI	SE	*k*	Total N	*Q*	I^2 (%)
Correlation, Total Score	.65***	.61-.69	.04	70	13,545	1,210.47***	93.78
Correlation, Factor 1	.42***	.36-.48	.04	46	7,478	344.79***	88.29
Correlation, Factor 2	.63***	.57-.69	.05	46	7,504	732.80***	94.00
Correlation, Facet 1	.30***	.24-.39	.04	29	6,761	184.52***	83.20
Correlation, Facet 2	.42***	.35-.48	.04	29	6,749	168.29***	83.80
Correlation, Facet 3	.59***	.51-.65	.06	30	6,766	341.12***	92.78
Correlation, Facet 4	.52***	.45-.59	.05	30	6,748	308.14***	90.54

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

Comparison of Studies Reporting a Protective Factors Measure

During data collection, risk measures which assessed protective factors rather than static and dynamic *risk* specific factors, such as the Structured Assessment of Protective Factors (SAPROF) were identified. These measures often have negative correlations with the risk they are developed to assess. Since these measures index somewhat different constructs than risk specific measures, which often correlate positively with risk estimates, the decision was made to analyze these measures separately.

After data collection and coding, the only measure featuring protective factors for which there was sufficient data for analysis was the SAPROF. The mean effect size for

the correlation between the PCL total score and SAPROF total score was *r*=-.50, which was statistically significant (*k*=7, 95% CI = -.618 – -.364; z = -6.354, *p* < .001). There was a considerable amount of heterogeneity in this sample, which is not explained by sampling error (*Q*(6)=40.77, *p*<.001). I^2 scores for the meta-analytic findings suggests 78.73% of the variance between the samples is not caused by sampling error and is better accounted for by between sample differences.

Although mean effect sizes were also calculated to assess for evidence of overlap between SAPROF risk measure total scores and PCL Factor and facet scores, the number of samples contributing to these calculations were very small, *k*=4 for Factor score correlations and *k*=2 for facet score correlations (Table 7), and thus were not conducive to further interpretation. Importantly, however, all effect sizes for all PCL Factor and facet scores with protective measure instruments demonstrated a *negative* relationship between the scores.

Table 7

Overall Meta-Analytic Effects and Measures of Heterogeneity, Correlations with the SAPROF Total Scores

	r	95% CI	SE	*k*	Total N	*Q*	I^2 (%)
Correlation, Total Score	-.50***	-.62 - -.36	.09	7	728	40.77***	78.73
Correlation, Factor 1	-.26**	-.46 - -.10	.09	4	467	11.30*	61.64
Correlation, Factor 2	-.35***	-.46 - -.24	.06	4	467	6.35	32.58
Correlation, Facet 1	-.10	-.20 - .01	.05	2	344	.86	0
Correlation, Facet 2	-.18	-.54 - .23	.21	2	345	19.99***	89.96
Correlation, Facet 3	-.32***	-.41 - -.22	.05	2	345	.82	0
Correlation, Facet 4	-.20***	-.30 - -.10	.05	2	345	1.26	.02

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

Comparison of Studies Based on Type of Risk Assessed

The next series of meta-analyses focused on exploring the mean effect sizes for correlations between PCL scores and risk assessment scores based on the type of recidivism measured by the risk assessment instrument (see Table 2). To explore this question, four separate meta-analyses were run for each PCL score, one featuring all studies for which the risk assessment instrument estimates general risk for recidivism, one for which the risk assessment instrument estimates violence recidivism, one for sexual recidivism instruments, and one for instruments estimating domestic violence risk. Overall effects and measures of heterogeneity are noted in Table 8. Of note, mean effect sizes were not calculated between PCL factor and facet scores and risk assessment scores for measures estimating domestic violence due to insufficient data. Only one study provided data for correlations with PCL facet scores, and none provided data for correlations with PCL factor scores.

Table 8

Overall Meta-Analytic Effects and Measure of Heterogeneity, Correlations Based on Type of Risk Assessed

	r	95% CI	SE	*k*	Total N	*Q*	I^2 (%)
Total Score							
General Risk	.63***	.53-.71	.08	23	3,717	404.75***	95.09
Violence Risk	.66***	.63-.69	.03	113	25,731	1,599.27***	93.46
Sexual Risk	.48***	.40-.55	.05	53	20,936	1,796.86***	97.82
Domestic Viol	.47***	.34-.58	.08	7	2,827	139.32***	93.33
Factor 1 Score							
General Risk	.35***	.27-.43	.05	13	2,207	55.94***	77.43
Violence Risk	.38***	.33-.42	.03	75	12,720	509.13***	86.29
Sexual Risk	.25***	.18-.33	.04	29	8,839	294.10***	91.69
Factor 2 Score							
General Risk	.72***	.65-.79	.08	13	2,206	128.78***	91.69
Violence Risk	.66***	.62-.70	.04	75	12,733	894.23***	93.16
Sexual Risk	.50***	.40-.59	.06	30	9,285	655.02***	97.09

	r	95% CI	SE	*k*	Total N	*Q*	I^2 (%)
Facet 1 Score							
General Risk	.24***	.13-.34	.06	9	1,516	42.55***	77.48
Violence Risk	.31***	.26-.36	.03	38	11,470	290.65***	85.96
Sexual Risk	.32***	.19-.43	.07	15	6,713	120.43***	89.74
Facet 2 Score							
General Risk	.47***	.30-.44	.04	7	1,338	15.85*	53.86
Violence Risk	.41***	.36-.46	.03	38	11,446	373.71***	85.69
Sexual Risk	.34***	.22-.46	.07	15	6,692	115.91***	90.51
Facet 3 Score							
General Risk	.58***	.48-.66	.07	7	1,338	46.04***	83.54
Violence Risk	.61***	.58-.65	.03	41	11,508	271.02***	81.41
Sexual Risk	.41***	.26-.55	.09	15	6,675	172.32***	94.35
Facet 4 Score							
General Risk	.59***	.47-.69	.09	7	1,338	48.69***	88.97
Violence Risk	.62***	.57-.66	.04	41	11,478	371.79***	91.40
Sexual Risk	.44***	.29-.57	.09	15	6,688	170.69***	93.81

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

Mean effects for PCL total, factor, and facet scores were statistically significant regardless of type of risk being assessed. Overall, a similar pattern in effect sizes was observed with moderate to large effect sizes between all types of risk measures and PCL total scores. At the factor and facet levels, effect sizes for Factor 2 scores were largest for all types of risk measures, with facet 3 and 4 effect sizes being moderate to large in size. Small to moderate effect sizes were noted for associations with Factor 1 and its constituent facets. When considering type of risk measured, larger effects were observed for associations with violence and general risk assessments than for sexual risk assessments for all scores except the facet 1 scores, for which the associations were larger for sexual and violence risk than for general risk. For all mean effects, a significant amount of variability among the included study effects was noted (Table 8). Indeed, between 81.41% and 97.82% of the variation observed was not explained by sampling error.

Comparison of Studies Reporting on Specific Risk Measures

Finally, meta-analyses were conducted for the correlation between PCL scores and any specific risk assessment measure for which there were at least five studies with information available regarding these correlations. Overall effects and measures of heterogeneity for correlations between PCL scores and the HCR-20, LSI-R, SAVRY, SORAG, Static-99, SVR-20, VRAG, and YLS/CMI are reported in Table 9. Mean effects for the correlation between PCL total scores and specific risk assessment measures were all statistically significant. Effect sizes were large (*r*s = .65-.78) for all risk measures except the Static-99 which demonstrated a moderate correlation with PCL total scores (*r* = .43). For all risk measures except the SAVRY, variance analyses suggested a significant amount of variability among the included study effects which is likely due to significant between study variability.

Although meta-analyses between PCL factor scores and specific risk assessment measures were conducted for all measures listed, several of the measures had few unique samples available for analysis (*k*=3 to *k*=4). As such, discussion will be limited to those measures for which there were more than 5 studies with reported Factor correlations; the HCR-20 and the VRAG. Mean effect sizes for both measures were statistically significant, and the correlations were in the moderate range for Factor 1 (*r*=.49 and *r*=.34 respectively) and in the large range for Factor 2 (*r*=.70 and *r*=.73 respectively). Similar to the patterns established above, effect sizes for Factor 2 scores are larger than those for Factor 1 scores. Variability statistics for both measures were statistically significant indicating large amounts of variability not accounted for by selection bias. This variability is most likely due to between study variables.

Finally, only two specific risk assessment measures had at least five reported correlations with PCL facet scores; the HCR-20 and the SAVRY. Mean effect sizes for both measures were statistically significant, and the correlations were in the moderate range for correlations with facet 1 (*r*=.30 and *r*=.33 respectively) and facet 2 (*r*=.44 and *r*=.47 respectively) scores, and in the large range for correlations with facet 3 (*r*=.63 and *r*=.64 respectively) and facet 4 (*r*=.55 and *r*=.64 respectively) scores. Similar to patterns established above, effect sizes for facet 3 and facet 4 are the largest, and the smallest effect sizes are observed for associations with facet 1 scores. Variability statistics for the correlations between PCL scores and HCR-20 scores were statistically significant for all facets indicating large amounts of variability not accounted for by selection bias. This variability is most likely due to between study variables. For the SAVRY, variability analyses suggested a significant amount of variability among the included study effects which is likely due to significant between study variability for facet 1 and facet 4 correlations. For facet 2 and facet 3 correlations, however, more modest amounts of variability were accounted for by between study variables.

Table 9

Overall Meta-Analytic Effects and Measure of Heterogeneity, Specific Risk Measures

	r	95% CI	SE	*k*	Total N	*Q*	I^2 (%)
Total Score							
HCR-20	.69***	.64-.73	.05	39	7,164	546.91***	91.80
LSI-R	.69***	.59-.77	.08	6	1,299	44.20***	88.41
SAVRY	.72***	.68-.75	.04	7	894	6.97	15.51
SORAG	.78***	.71-.84	.08	5	1,920	33.60***	91.31
Static-99	.43***	.33-.52	.06	10	5,477	183.01***	94.42
SVR-20	.67***	.59-.74	.07	9	2,067	68.25***	87.04
VRAG	.69***	.64-.73	.04	23	7,271	201.81***	90.27
YLS/CMI	.65***	.54-.74	.03	7	1,037	57.22***	87.35
Factor 1 Score							
HCR-20	.49***	.44-.53	.03	30	4,440	137.21***	74.61

	r	95% CI	SE	*k*	Total N	*Q*	I^2 (%)
LSI-R	.29***	.21-.36	.04	4	884	6.27	36.44
SAVRY	.45***	.37-.53	.05	3	455	3.67	10.82
SORAG	.37***	.28-.45	.05	4	1,171	9.98*	55.30
Static-99	.19***	.14-.24	.03	4	1,452	4.32	.20
SVR-20	.37***	.12-.57	.13	4	1,160	64.57***	94.06
VRAG	.34***	.26-.41	.04	14	2,729	55.17***	75.43
YLS/CMI	.40***	.30-.49	.02	4	605	7.38	45.91
Factor 2 Score							
HCR-20	.70***	.65-.74	.04	30	4,462	223.48***	86.28
LSI-R	.72***	.53-.84	.16	4	884	50.00***	95.09
SAVRY	.74***	.69-.79	.06	3	455	4.81	39.10
SORAG	.78***	.71-.83	.08	4	1,171	17.83***	81.35
Static-99	.52***	.39-.63	.08	4	1,453	26.80***	89.30
SVR-20	.60***	.53-.65	.05	4	1,160	8.24*	48.70
VRAG	.73***	.66-.78	.07	14	2,718	129.34***	90.13
YLS/CMI	.65***	.50-.76	.03	4	604	32.51***	86.41
Facet 1 Score							
HCR-20	.30***	.20-.39	.05	14	3,692	113.57***	85.94
LSI-R	.16	.00-.31	.08	4	911	19.17***	82.48
SAVRY	.33***	.22-.44	.06	5	689	12.87**	61.41
SVR-20	.51***	.43-.57	.05	4	1,822	1.205	0.00
VRAG	.29***	.18-.40	.06	5	2,388	17.87**	77.69
YLS/CMI	.28***	.13-.41	.03	3	455	7.96**	61.73
Facet 2 Score							
HCR-20	.44***	.36-.51	.05	14	3,687	82.78***	81.54
LSI-R	.34***	.24-.43	.06	3	822	7.48*	59.89
SAVRY	.47***	.41-.53	.04	5	689	.30	.00
SVR-20	.50***	.43-.57	.05	4	1,815	3.32	.00
VRAG	.31***	.16-.44	.08	5	2,383	34.16***	87.31
YLS/CMI	.38***	.28-.46	.03	3	455	3.50	20.04
Facet 3 Score							
HCR-20	.63***	.56-.69	.06	14	3,676	102.83***	88.10
LSI-R	.57***	.40-.70	.11	3	822	29.15***	90.05
SAVRY	.64***	.60-.69	.04	5	689	2.68	.00
SVR-20	.58***	.51-.64	.05	4	1,809	.97	.00
VRAG	.61***	.50-.71	.08	5	2,371	57.17***	87.94
YLS/CMI	.58***	.42-.70	.03	3	455	15.99***	79.52
Facet 4 Score							
HCR-20	.55***	.48-.62	.05	14	3,659	100.68***	83.91
LSI-R	.51***	.43-.59	.06	3	822	7.538*	60.67
SAVRY	.64***	.57-.71	.06	5	689	13.776**	60.64
SVR-20	.53***	.46-.60	.05	4	1,812	4.15	.02
VRAG	.72***	.65-.78	.07	5	2,354	44.36***	84.13
YLS/CMI	.57***	.40-.71	.04	3	455	19.95***	83.19

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

Publication Bias

Analyses were conducted to assess for inflated effects due to the under-publication of nonsignificant effect sizes and the inclusion of only published studies and publicly available theses and dissertations. First, Egger's regression tests for each meta-analytic result were conducted. Overall, the Egger's test results suggested most (80.4%) of the effect sizes fell within the expected distribution as described by the funnel plots for each meta-analysis providing evidence against the presence of publication bias (Tables 10-14). No systematic pattern of potential publication bias emerged. Specifically, Egger's tests for the correlation between all risk measures and all PCL Factor 1 scores, PCL-R total, Factor 1, facet 1, and facet 2 scores, and PCL-YV total and facet 2 scores were significant suggesting potential publication bias. Additionally, when analyses were divided by risk type measures, correlations with PCL total and facet 4 scores were significant for actuarial measures as were correlations with PCL factor 1 scores for violence risk measures and PCL factor 2 scores for general risk measures. Finally, when considering specific risk assessment measures, Egger's tests between LSI-R scores and PCL total, facet 1 and facet 3 scores were significant as were those for correlations between the SORAG and PCL total and Factor 2 scores, VRAG and Factor 1 scores, Static-99 and Factor 2 scores, YLS/CMI Factor 2, facet 3, and facet 4 scores and the SAVRY and facet 4 scores.

To further assess for publication bias Duval and Tweedie's trim-and-fill procedure was completed for each correlation. The number of studies imputed using this method as well as the adjusted correlation and difference between the adjusted correlation and observed correlations are in Tables 10-14. Overall, nearly all correlations (97.3%)

demonstrated only small differences between adjusted and observed correlations, providing further evidence for lack of publication bias. Only the trim-and-fill procedures for the correlation between general risk measures and PCL total scores, sexual risk measures and facet 2 scores, and sexual risk measures and facet 3 scores resulted in large differences between adjusted and observed effect sizes (differences of .14, .15, and .13 respectively) suggesting funnel plot asymmetry and potential publication bias.

Finally, Rosenthal's fail-safe *N* was calculated for each meta-analytic effect (Tables 10-14). These calculations suggested, for most correlations, large numbers (80-1,281,000, median 9,107) of studies with null effects would have to be published to render the current observed effect sizes non-significant ($p>.05$) for analyses consisting of more than 5 samples. Given the very large values suggested by Rosenthal's fail-safe *N* calculations and the unlikelihood that this many unpublished studies exist with null findings, it is likely that our current findings are not significantly influenced by selective publishing practices.

Given the sum of data on potential publication bias, it is unlikely that the effects described in the present study were unduly impacted by publication bias or the inclusion of only publicly available studies. It is, however, important to note that Rosenthal's fail-safe *N* suggested that only 10 articles with null results regarding the correlation between PCL total scores and SAPROF total scores would have to be published to render the mean effect of the current analyses non-significant ($p<.05$). Due to the small number of studies included in the calculation of this effect size, however, results should likely be interpreted cautiously.

Table 10

Publication Bias Analyses for All Risk Assessment Measures

			Adjusted		Difference	
	Z	Imputed	Effect Size	95% CI		Fail-Safe N
Total Score	1.67	34	.52	.49-.56	.07	1,056,000***
Factor 1	.91	23	.26	.23-.31	.08	78,365***
Factor 2	3.23***	22	.55	.51-.58	.08	369,282***
Facet 1	1.42	10	.22	.17-.27	.05	16,022***
Facet 2	.77	8	.33	.28-.38	.04	28,082***
Facet 3	.32	12	.51	.45-.55	.05	89,120***
Facet 4	-1.19	0	.59	.53-.62	.00	106,660***

Note. First column provides Z value from the Egger's regression test. Imputed column notes the numbers of filler studies added during Duval and Tweedie's trim-and-fill analysis. Adjusted effects sizes are calculated utilizing imputed articles from Duval and Tweedie's trim-and-fill analysis. Diff = the difference between he adjusted and observed effect sizes. Fail-safe N is the number of articles with null findings that would need to be published to render the observed effect nonsignificant.
*p<.05. **p<.01. ***p<.001.

Table 11

Publication Bias Analyses for Correlations Based on PCL Version Utilized

			Adjusted			
	Z	Imputed	Effect Size	95% CI	Diff	Fail-Safe N
PCL-R						
Total Score	2.67**	18	.55	.51-.59	.05	1,281,000***
Factor 1 Score	3.80***	15	.25	.21-.31	.07	81,031***
Factor 2 Score	1.92	10	.59	.54-.64	.05	419,674***
Facet 1 Score	3.24***	12	.19	.12-.24	.08	8,441***
Facet 2 Score	2.78**	6	.29	.24-.35	.06	12,728***
Facet 3 Score	1.82	6	.52	.46-.57	.04	55,702***
Facet 4 Score	-.15	2	.60	.54-.64	.01	73,476***
PCL:SV						
Total Score	1.16	7	.48	.37-.57	.08	19,077***
Factor 1 Score	.23	5	.29	.19-.39	.07	4,188***
Factor 2 Score	-.22	0	.55	.44-.64	0	12,210***
Facet 1 Score	-1.82	0	.18	.00-.34	0	80***
Facet 2 Score	-1.04	0	.34	.13-.52	0	312***
Facet 3 Score	-.62	0	.42	.12-.64	0	478***
Facet 4 Score	-.75	0	.35	.11-.55	0	323***
PCL:YV						
Total Score	-2.53*	0	.56	.43-.67	0	12,644***
Factor 1 Score	-1.53	0	.42	.36-.48	0	688***

			Adjusted			
	Z	Imputed	Effect Size	95% CI	Diff	Fail-Safe N
Factor 2 Score	-2.52*	0	.68	.59-.75	0	2,377***
Facet 1 Score	-1.73	0	.33	.22-.42	0	850***
Facet 2 Score	-1.20	1	.49	.44-.53	.02	1,822***
Facet 3 Score	-1.04	0	.62	.57-.66	0	3,622***
Facet 4 Score	-2.17	0	.64	.58-.70	0	4,159***

Note. First column provides Z value from the Egger's regression test. Imputed column notes the numbers of filler studies added during Duval and Tweedie's trim-and-fill analysis. Adjusted effects sizes are calculated utilizing imputed articles from Duval and Tweedie's trim-and-fill analysis. Diff = the difference between he adjusted and observed effect sizes. Fail-safe N is the number of articles with null findings that would need to be published to render the observed effect nonsignificant.
*p<.05. **p<.01. ***p<.001.

Table 12

Publication Bias Analyses for Correlations, Actuarial Risk Assessment Measures and Structured Professional Judgment Assessment Measures

			Adjusted			
	Z	Imputed	Effect Size	95% CI	Diff	Fail-Safe N
Actuarial						
Total Score	2.84**	0	.28	.24-.32	.00	17,991***
Factor 1 Score	1.55	5	.60	.54-.65	.03	126,232***
Factor 2 Score	1.11	0	.25	.19-31	.00	4,079***
Facet 1 Score	1.47	0	.33	.27-.39	.00	6,242***
Facet 2 Score	.49	1	.53	.48-.58	.01	24,379***
Facet 3 Score	.32	0	.63	.57-.68	.00	39,121***
Facet 4 Score	2.84**	0	.28	.24-.32	.00	17,991***
SPJ						
Total Score	1.77	19	.58	.52-.62	.07	162,176***
Factor 1 Score	1.20	13	.35	.27-.41	.07	20,820***
Factor 2 Score	-.09	0	.63	.57-.69	.00	58,306***
Facet 1 Score	1.20	5	.25	.19-.33	.05	4,391***
Facet 2 Score	.80	8	.36	.30-.42	.06	8,609***
Facet 3 Score	.38	9	.52	.43-.58	.07	20,874***
Facet 4 Score	-1.03	0	.52	.45-.59	.00	16,565***

Note: First column provides Z value from the Egger's regression test. Imputed column notes the numbers of filler studies added during Duval and Tweedie's trim-and-fill analysis. Adjusted effects sizes are calculated utilizing imputed articles from Duval and Tweedie's trim-and-fill analysis. Diff = the difference between he adjusted and observed effect sizes. Fail-safe N is the number of articles with null findings that would need to be published to render the observed effect nonsignificant.
*p<.05. **p<.01. ***p<.001.

Table 13

Publication Bias Analyses for Correlations Based on Type of Risk Assessed

			Adjusted			
	Z	Imputed	Effect Size	95% CI	Diff	Fail-Safe N
Total Score						
General Risk	.23	9	.49	.37-.60	.14	15,730***
Violence Risk	.60	20	.62	.58-.65	.04	513,651***
Sexual Risk	-.04	0	.48	.40-.55	.00	83,469***
Domestic Viol	-1.13	2	.52	.41-.62	.05	1,753***
Factor 1 Score						
General Risk	1.21	1	.34	.25-.41	.01	1,284***
Violence Risk	2.06*	15	.31	.26-.36	.07	44,144***
Sexual Risk	1.13	5	.20	.13-.27	.05	4,764***
Factor 2 Score						
General Risk	1.79	0	.72	.65-.79	.00	7,919***
Violence Risk	-.57	0	.66	.62-.70	.00	186,092***
Sexual Risk	.71	0	.50	.40-.59	.00	24,033***
Facet 1 Score						
General Risk	2.68**	1	.22	.11-.32	.02	223***
Violence Risk	1.11	1	.30	.24-.35	.01	9,605***
Sexual Risk	.35	2	.23	.09-.35	.09	1,120***
Facet 2 Score						
General Risk	1.13	0	.47	.30-.44	.00	482***
Violence Risk	.83	3	.40	.35-.45	.01	16,433***
Sexual Risk	-.40	4	.19	.05-.33	.15	1,331***
Facet 3 Score						
General Risk	-.07	0	.58	.48-.66	.00	1,417***
Violence Risk	.31	6	.59	.55-.62	.02	51,956***
Sexual Risk	.35	5	.28	.14-.41	.13	1,950***
Facet 4 Score						
General Risk	1.64	0	.59	.47-.69	.00	1,366***
Violence Risk	-1.92	6	.65	.60-.69	.03	58,315***
Sexual Risk	-.46	0	.44	.29-.57	.00	2,360***

Note. First column provides Z value from the Egger's regression test. Imputed column notes the numbers of filler studies added during Duval and Tweedie's trim-and-fill analysis. Adjusted effects sizes are calculated utilizing imputed articles from Duval and Tweedie's trim-and-fill analysis. Diff = the difference between he adjusted and observed effect sizes. Fail-safe N is the number of articles with null findings that would need to be published to render the observed effect nonsignificant.
*p<.05. **p<.01. ***p<.001.

Table 14

Publication Bias Calculations for the Mean Effects, Specific Risk Assessment Measures

			Adjusted			
	Z	Imputed	Effect Size	95% CI	Diff	Fail-Safe N
Total Score						
HCR-20	1.22	2	.68	.62-.72	.01	55,748***
LSI-R	2.51*	0	.69	.59-.77	.00	1,822***
SAVRY	-1.84	2	.74	.71-.77	.02	1,805***
SORAG	2.42*	0	.78	.71-.84	.00	3,164***
Static-99	1.01	0	.43	.33-.52	.00	3,083***
SVR-20	-.69	0	.67	.59-.74	.00	3,754***
VRAG	-.14	0	.69	.64-.73	.00	34,876***
YLS/CMI	-1.81	0	.65	.54-.74	.00	1,605***
Factor 1 Score						
HCR-20	1.84	7	.44	.38-.49	.05	11,018***
LSI-R	1.09	1	.27	.20-.35	.02	104***
SAVRY	-.31	0	.45	.37-.53	.00	112***
SORAG	.88	1	.35	.25-.43	.02	221***
Static-99	1.71	1	.17	.12-.22	.02	76***
SVR-20	.13	0	.37	.12-.57	.00	238***
VRAG	2.46**	0	.34	.26-.41	.00	1,414***
YLS/CMI	-1.88	0	.40	.30-.49	.00	147***
Factor 2 Score						
HCR-20	-.46	6	.74	.69-.77	.04	30,379***
LSI-R	2.59	0	.72	.53-.84	.00	856***
SAVRY	-1.96	0	.74	.69-.79	.00	440***
SORAG	2.65**	1	.75	.66-.82	.03	1,541***
Static-99	3.11**	0	.52	.39-.63	.00	624***
SVR-20	.23	1	.58	.53-.64	.02	689***
VRAG	.73	3	.70	.62-.75	.03	10,016***
YLS/CMI	-3.20***	0	.65	.50-.76	.00	531***
Facet 1 Score						
HCR-20	1.17	0	.30	.20-.39	.00	1,335***
LSI-R	3.39***	0	.16	.00-.31	.00	22***
SAVRY	.08	1	.31	.21-.41	.02	142***
SVR-20	-.50	0	.51	.43-.57	.00	129***
VRAG	1.22	2	.23	.09-.35	.06	249***
YLS/CMI	-.84	0	.28	.13-.41	.00	37***
Facet 2 Score						
HCR-20	.75	2	.41	.33-.49	.03	2,938***
LSI-R	.19	0	.34	.24-.43	.00	110***
SAVRY	.24	0	.47	.41-.53	.00	317***
SVR-20	-1.43	1	.52	.45-.58	.02	113***
VRAG	1.08	0	.31	.16-.44	.00	247***

			Adjusted			
	Z	Imputed	Effect Size	95% CI	Diff	Fail-Safe N
YLS/CMI	-1.80	0	.38	.28-.46	.00	70***
Facet 3 Score						
HCR-20	.39	2	.60	.52-.68	.03	7,688***
LSI-R	-5.23***	0	.57	.40-.70	.00	376***
SAVRY	-.81	0	.64	.60-.69	.00	699***
SVR-20	.03	0	.58	.51-.64	.00	194***
VRAG	.60	1	.59	.49-.68	.02	1,514***
YLS/CMI	-2.01*	0	.58	.42-.70	.00	213***
Facet 4 Score						
HCR-20	-1.80	1	.56	.49-.63	.01	5,790***
LSI-R	-1.03	1	.55	.45-.64	.04	285***
SAVRY	-3.08**	1	.66	.59-.72	.02	718***
SVR-20	-.96	2	.55	.49-.61	.02	140***
VRAG	-.54	0	.72	.65-.78	.00	2,565***
YLS/CMI	-3.14**	0	.57	.40-.71	.00	215***

Note. First column provides Z value from the Egger's regression test. Imputed column notes the numbers of filler studies added during Duval and Tweedie's trim-and-fill analysis. Adjusted effects sizes are calculated utilizing imputed articles from Duval and Tweedie's trim-and-fill analysis. Diff = the difference between he adjusted and observed effect sizes. Fail-safe N is the number of articles with null findings that would need to be published to render the observed effect nonsignificant.
*p<.05. **p<.01. ***p<.001.

Moderator Analyses

All moderator analyses were conducted on the complete set of studies featuring risk assessment measures, regardless of PCL measure utilized. Moderator analyses were completed for associations with PCL total, Factor, and facet scores separately within this dataset. The only data points not considered in these analyses were those for which the SAPROF was the assessment measure due to the use of protective factors rather than risk factors in the coding of this measure. Additionally, categories for which less than five samples were included (e.g., only one sample reported data for offenders on conditional release) were not retained in moderator analyses.

Moderation analyses revealed no significant differences in the correlation between PCL total scores and risk measure total scores as a function of (a) sex

composition of the sample, (b) method of scoring the PCL measure (e.g., file and interview or file only), (c) type of publication in which data were reported, (d) race composition of the sample, measured in percentage of the population Caucasian, (e) sample type (correctional, institutional, or probation/parole), (f) age composition of the sample (i.e., adult or juvenile sample and mean age), (g) reliability as measured by the Cronbach's alpha, or (h) year of publication of the article (Table 14 and Table 15). However, a significant difference was found in the effect size based on the country or continent in which the study was conducted (Q = 7.558, p = .023). Effect sizes for PCL total scores were larger for samples from European countries (r = .61) and Canada (r = .60) than for samples from the United States (r = .48). For this calculation studies conducted in European countries were combined into a single group as individual countries often did not have enough publications to constitute their own group for moderation analyses.

Moderation analyses revealed no significant differences in the correlation between PCL Factor 1 scores and risk measure total scores for any of the moderators listed. Additionally, no significant differences in the correlations between PCL Factor 2 scores and risk measure total scores were observed as a function of (a) sex composition of the sample, (b) method of scoring the PCL measure (e.g., file and interview or file only), (c) type of publication in which data were reported, (d) race composition of the sample, (e) country of study, (f) age composition of the sample (i.e., adult or juvenile sample), or (g) year of publication of the article. In contrast, a significant difference was found in the correlation between PCL Factor 2 scores and risk measure total scores based on the sample type (Q = 11.483, p = .003), age when calculated as mean sample age (Q =

4.886, p = .027) and reliability coefficient, measured by Cronbach's alpha, for PCL Factor 2 scores (Q = 8.492, p = .004). Specifically, effect sizes were larger for probation or parole samples (r = .74) than for correctional samples (r = .66). Both probation or parole samples and correctional samples had larger effect sizes than for samples using institutionalized participants (r = .54). Meta-analytic effects were also larger as the reliability coefficient, Cronbach's alpha, increased (unstandardized b = 2.225) and as mean age of the sample decreased (unstandardized b = -.011).

Finally, moderation analyses revealed significant differences in the correlation between PCL facet 2 and facet 4 scores and risk measure total scores based on age, calculated as the mean age of the sample (Q = 5.905, p = .015; and Q = 8.434, p = .004 respectively), and sample type (Q = 6.965, p = .031 and Q = 8.582, p = .014 respectively). Additionally, a significant difference was observed between facet 2 scores and risk measure total scores based on the reliability coefficient, measured by Cronbach's alpha (Q = 4.429, p = .035). All moderation effects were in the same direction as mentioned above (see tables 14 and 15 for more detail)

Table 15

Moderator Analyses for Continuous Variables

	Total	Factor 1	Factor 2	Facet 1	Facet 2	Facet 3	Facet 4
Age (Mean)							
Number of Effects	169	98	99	54	52	55	55
Unstandardized coefficient b_1	.0004	-.004	-.011	-.004	-.008	-.006	-0.012
Standard error of b_1	.003	.003	.005	.003	.003	.004	.004
Test of $b_1 \neq 0$	Z = -.140, p =.888	Z = -1.164, p =.244	Z = -2.210, p = .027	Z = -1.120, p = .263	Z = -2.430, p = .015	Z = -1.331, p = .183	Z = -2.904, p = .004
Test of variability not explained by moderator	Q_w[167] = 4121.059, p <.001	Q_w[96] = 785.185, p <.001	Q_w[97] = 1682.557, p <.001	Q_w[52] = 444.067, p <.001	Q_w[50] = 516.930, p <.001	Q_w[53] = 595.750, p <.001	Q_w[53] = 641.543, p <.001
Race							
Number of Effects	99	52	52	28	26	29	29
Unstandardized coefficient b_1	.003	-.001	.001	.002	-.002	.002	.002
Standard error of b_1	.002	.002	.002	.002	.003	.003	.004
Test of $b_1 \neq 0$	Z = 1.550, p = .121	Z = -.420, p = .674	Z = .590, p = .555	Z = .789, p = .430	Z = -.875, p = .382	Z = .523, p = .601	Z = .629, p = .529
Test of variability not explained by moderator	Q_w[97] = 2651.230, p <.001	Q_w[50] = 467.866, p <.001	Q_w[50] = 914.365, p <.001	Q_w[26] = 346.877, p <.001	Q_w[24] = 370.202, p <.001	Q_w[27] = 467.459, p <.001	Q_w[27] = 481.658, p <.001
Year of Publication							
Number of Effects	181	107	108	58	56	59	59
Unstandardized coefficient b_1	.004	.003	-.003	-.002	.001	.007	-.008
Standard error of b_1	.005	.004	.005	.006	.006	.007	.008

	Total	Factor 1	Factor 2	Facet 1	Facet 2	Facet 3	Facet 4
Test of $b_1 \neq 0$	Z = .817, p = .414	Z = .804, p = .421	Z = -.502, p = .616	Z = -.428, p = .669	Z = .169, p = .866	Z = .922, p = .357	Z = -1.044, p = .297
Test of variability not explained by moderator	Q_w[179] = 4542.316, p <.001	Q_w[105] = 830.196, p <.001	Q_w[106] = 1948.606, p <.001	Q_w[56] = 490.726, p <.001	Q_w[54] = 583.479, p <.001	Q_w[57] = 654.333, p <.001	Q_w[57] = 695.867, p <.001
Reliability Coefficient							
Number of Effects	25	18	18	8	8	8	8
Unstandardized coefficient b_1	1.389	-.039	2.225	.324	-.998	1.863	.138
Standard error of b_1	.817	.305	.763	.562	.474	2.916	1.372
Test of $b_1 \neq 0$	Z = 1.701, p = .089	Z = -.128, p = .898	Z = 2.914, p = .004	Z = .576, p = .564	Z = -2.104, p = .035	Z = .639, p = .523	Z = .101, p = .920
Test of variability not explained by moderator	Q_w[23] = 392.716, p <.001	Q_w[16] = 87.179, p <.001	Q_w[16] = 223.345, p <.001	Q_w[5] = 55.062, p <.001	Q_w[5] = 39.384, p <.001	Q_w[5] = 125.850, p <.001	Q_w[5] = 96.412, p <.001

Table 16

Moderator Analyses for Categorical Variables

	Total	Factor 1	Factor 2	Facet 1	Facet 2	Facet 3	Facet 4
Country							
United States							
Number of Effects	32	19	19	7	7	7	7
Number of Participants	12,258	5,918	5,920	1,205	1,205	1,205	1,205
Mean *r*	.485	.327	.557	.297	.368	.473	.545
Canada							
Number of Effects	69	46	47	14	12	15	15
Number of Participants	15,354	8,488	8,936	1,910	1,732	1,831	1,829
Mean *r*	.614	.389	.642	.278	.326	.587	.580
European Countries							
Number of Effects	76	38	38	33	33	33	33
Number of Participants	19,629	5,714	5,722	13,978	13,940	13,892	13870
Mean *r*	.600	.297	.629	.243	.293	.497	.572
Test of variability explained by moderator	$Q_b[2] = 7.558$, $p = .023$	$Q_b[2] = 4.756$, $p = .093$	$Q_b[2] = 2.128$, $p = .345$	$Q_b[2] = .489$, $p = .783$	$Q_b[2] = .273$, $p = .872$	$Q_b[2] = 3.934$, $p = .140$	$Q_b[2] = .206$, $p = .902$
Test of variability not explained by moderator	$Q_w[174] = 4076.107$, $p < .001$	$Q_w[100] = 782.738$, $p < .001$	$Q_w[101] = 1820.483$, $p < .001$	$Q_w[51] = 464.174$, $p < .001$	$Q_w[49] = 544.042$, $p < .001$	$Q_w[52] = 573.596$, $p < .001$	$Q_w[52] = 640.829$, $p < .001$
Age							
Adult							
Number of Effects	108	66	67	31	29	29	29

	Total	Factor 1	Factor 2	Facet 1	Facet 2	Facet 3	Facet 4
Number of Participants	33,025	10,842	11,314	13,025	12,805	12,769	12,782
Mean *r*	.597	.349	.649	.309	.387	.584	.604
Juvenile							
Number of Effects	21	6	6	8	8	8	8
Number of Participants	2,636	777	776	1,008	1,008	1,008	1,008
Mean *r*	.523	.395	.621	.265	.465	.589	.600
Test of variability explained by moderator	$Q_b[1] = 1.725$, $p = .189$	$Q_b[1] = .371$, $p = .542$	$Q_b[1] = .109$, $p = .741$	$Q_b[1] = .430$, $p = .512$	$Q_b[1] = 1.816$, $p = .178$	$Q_b[1] = .010$, $p = .919$	$Q_b[1] = .006$, $p = .937$
Test of variability not explained by moderator	$Q_w[122] = 3481.593$, $p < .001$	$Q_w[65] = 405.627$, $p < .001$	$Q_w[66] = 1124.939$, $p < .001$	$Q_w[33] = 323.736$, $p < .001$	$Q_w[31] = 372.424$, $p < .001$	$Q_w[31] = 319.012$, $p < .001$	$Q_w[31] = 332.017$, $p < .001$
Sample Type							
Correctional							
Number of Effects	89	43	44	30	28	28	28
Number of Participants	28,218	8,748	9,189	14,385	14,165	14,129	14,147
Mean *r*	.574	.308	.656	.263	.384	.572	.611
Institutional							
Number of Effects	59	45	45	17	17	20	20
Number of Participants	10,096	8,288	8,302	2,017	2,021	2,108	2,066
Mean *r*	.581	.335	.542	.261	.284	.514	.485
Probation/Parole							
Number of Effects	11	9	9	3	3	3	3
Number of Participants	1,802	1,592	1,592	464	464	464	464
Mean *r*	.687	.376	.757	.256	.509	.702	.713

	Total	Factor 1	Factor 2	Facet 1	Facet 2	Facet 3	Facet 4
Test of variability explained by moderator	$Q_b[2]$ = 3.108, p = .211	$Q_b[2]$ = 1.162, p = .559	$Q_b[2]$ = 14.104, p < .001	$Q_b[2]$ = .004, p = .998	$Q_b[2]$ = 6.965, p = .031	$Q_b[2]$ = 4.190, p = .123	$Q_b[2]$ = 8.582, p = .014
Test of variability not explained by moderator	$Q_w[156]$ = 4088.704, p <.001	$Q_w[94]$ = 709.305, p <.001	$Q_w[95]$ = 1656.469, p <.001	$Q_w[47]$ = 457.389, p <.001	$Q_w[45]$ = 480.119, p <.001	$Q_w[48]$ = 565.640, p <.001	$Q_w[48]$ = 592.941, p <.001
Sex							
Male							
Number of Effects	131	76	77				
Number of Participants	34,242	16,179	16,613				
Mean r	.562	.312	.622				
Female							
Number of Effects	6	7	7				
Number of Participants	584	651	651				
Mean r	.534	.324	.515				
Test of variability explained by moderator	$Q_b[1]$ = .071, p = .789	$Q_b[1]$ = .025, p = .876	$Q_b[1]$ = 1.383, p = .240				
Test of variability not explained by moderator	$Q_w[135]$ = 3486.089, p <.001	$Q_w[81]$ = 556.524, p <.001	$Q_w[82]$ = 1592.701, p <.001				
Type of Document							
Journal Article							
Number of Effects	158	88	89				
Number of Participants	40,954	18,480	18,939				
Mean r	.577	.329	.609				
Masters Thesis							

	Total	Factor 1	Factor 2	Facet 1	Facet 2	Facet 3	Facet 4
Number of Effects	4	6	6				
Number of Participants	307	384	384				
Mean *r*	.774	.448	.697				
Doctoral Dissertation							
Number of Effects	19	13	13				
Number of Participants	6,518	1,794	1,793				
Mean *r*	.634	.359	.700				
Test of variability explained by moderator	$Q_b[2] = 5.623$, $p = .060$	$Q_b[2] = 2.000$, $p = .368$	$Q_b[2] = 3.416$, $p = .181$				
Test of variability not explained by moderator	$Q_w[178] = 4361.189$, $p < .001$	$Q_w[104] = 826.179$, $p < .001$	$Q_w[105] = 1887.557$, $p < .001$				
Information Used to Assess the PCL							
File and Interview							
Number of Effects	63	35	35	25	25	25	25
Number of Participants	19,213	7,474	7,917	13,175	13,133	13,097	13,115
Mean *r*	.566	.364	.625	.292	.396	.589	.619
File Only							
Number of Effects	64	39	39	15	15	18	18
Number of Participants	14,785	8,658	8,686	2,085	2,085	2,184	2,182
Mean *r*	.602	.309	.604	.266	.363	.516	.543
Test of variability explained by moderator	$Q_b[1] = .753$, $p = .386$	$Q_b[1] = 1.441$, $p = .230$	$Q_b[1] = .174$, $p = .677$	$Q_b[1] = .158$, $p = .691$	$Q_b[1] = .326$, $p = .568$	$Q_b[1] = 1.418$, $p = .234$	$Q_b[1] = 1.669$, $p = .196$

	Total	Factor 1	Factor 2	Facet 1	Facet 2	Facet 3	Facet 4
Test of variability not explained by moderator	$Q_w[121]$ = 3130.287, p <.001	$Q_w[68]$ = 651.474, p <.001	$Q_w[69]$ = 1594.764, p <.001	$Q_w[34]$ = 420.231, p <.001	$Q_w[34]$ = 456.223, p <.001	$Q_w[37]$ = 502.070, p <.001	$Q_w[37]$ = 465.624, p <.001

CHAPTER IV

Discussion

Meta-Analytic Findings

The current study was completed to provide information to practitioners in the field of forensic psychology concerning the potential overlap between scores of PCL measures and established risk assessment measures. This information will aid practitioners in determining the weight given to each piece of evidence, especially when combining multiple measures and sources of information to obtain their final judgment on an individual's risk of violence or recidivism. In total, I utilized information from 87 samples which provided 187 entries pertaining to correlations between risk measures and PCL measures. Of those, most (96.8%) contained a correlation between the PCL total scores and risk measure total score, with much fewer reporting results for PCL factors or facets. Approximately one half reported correlations with PCL factor scores (57.8% for factor 1 and 58.2% for factor 2) and approximately one third reported correlations for PCL facet scores (31.6%, 30.4%, 32.1%, and 32.1% for facets 1-4 respectively).

Overall, findings indicate a large correlation between risk assessment measures and PCL total scores (r=.59). Further, large correlations were observed between risk assessment measure total scores and PCL Factor 2 (r = .63) scores and its constituent facets, facet 3 (r = .56) and facet 4 (r = .59). By contrast, moderate correlations were observed between risk measure total scores and Factor 1 scores (r = .34) and facet 2 scores (r = .37) and the smallest correlation was observed with facet 1 scores (r = .27). This is not wholly unsurprising given prior research suggesting Factor 2 scores correlate more strongly with risk measures than Factor 1 scores (see for example Neal et al, 2015;

Simourd & Hoge, 2000; Warrne et al, 2005). Since Factor 2 scores are composed of the facet 3 and facet 4 scores, it also stands to reason that these should correlate more strongly with risk measure scores than the facets which compose the Factor 1 score (i.e., facet 1 and facet 2). Importantly, both facets 3 and 4 displayed large effects of similar magnitude with overall risk, which indicates that both the historical content captured in facet 4 (i.e., juvenile and adult antisocial behavior) as well as impulsive, irresponsible personality traits indexed by facet 3 are strongly relevant to determining risk. While Factor 1 scores showed more modest associations with risk measures, these effects were statistically significant and robust across all analyses. Likewise effect sizes were consistently larger for facet 2 than facet 1, suggesting callous, remorseless traits indexed by the affective facet of the PCL (facet 2) may be more relevant to determining risk than the interpersonal features of psychopathy (facet 1). Overall, these results indicate that externalizing proneness (as measured by Factor 2) overlaps substantially with what is assessed by risk measures; however, core dynamic personality traits underlying psychopathy may confer additional risk-relevant information, albeit to a lesser extent.

Similar results were observed when subdividing the meta-analysis to consider the version of the PCL utilized. Indeed, total, Factor 2, facet 3 and facet 4 correlations with risk scores were larger for all individual PCL measures than Factor 1, facet 1, and facet 2 scores, which fell into the small to moderate categories for each measure. Notably, however, most of the PCL:SV effect sizes were smaller than those for the PCL-R and PCL:YV. Given that the PCL:SV was designed as a screening measure, it is possible that the associations are weaker for this measure due to the smaller number of items contributing to factor and facet scores, reducing their reliability. Further, it is plausible

that the additional items on the PCL-R and PCL:YV that are not contained within the screening form of the PCL index risk-relevant content, augmenting their observed effect sizes.

Actuarial versus Structured Professional Judgement Risk Measures

The next set of analyses considered divisions of the data based on risk measure type (i.e., actuarial or SPJ) and type of risk assessed by the measure (i.e., general risk, violence risk, sexual risk, or domestic violence risk). Overall, for both types of risk assessment measures, the PCL total scores demonstrated moderate to large effect sizes (.47-.66). Factor 2, facet 3, and facet 4 scores demonstrated large effect sizes, and Factor 1, facet 1, and facet 2 scores evidenced small to moderate effect sizes with facet 1 scores demonstrating the lowest effect sizes across all meta-analyses. It is notable that effect sizes between PCL scores and structured professional judgment measures tended to be larger than effect sizes for associations between PCL scores and actuarial measures. Structured professional judgment measures require more subjective coding rather than the more structured actuarial instruments. This suggests that actuarial measures may miss out on personality characteristics that may be related to risk of reoffending due to the strict coding schemas and the disallowance of clinicians to adjust risk based on clinical experience with personality disorders.

General, Violent, and Sexual Risk

Similar patterns were observed when accounting for type of risk measured (i.e., general, violent, or sexual). Large effect sizes were observed for associations with PCL total, Factor 2, facet 3, and facet 4 scores for measures of general and violence risk. Moderate effect sizes were observed for associations with PCL Factor 1 and facet 2

scores and the smallest effect sizes were observed for associations with facet 1 scores. The findings of the current study are in line with previous research which reports wide ranges of predictive capabilities of the PCL-R with respect to recidivism, but appear to suggest these measures may correlate with violence and general risk more strongly than sexual risk (for example see Barbaree et al, 2001; Caperton, 2005; Larsen et al, 2020; Sjostedt & Langstrom, 2002; Walters, 2003). Similarly, our results demonstrate PCL measures, Factors, and facets correlate more strongly with measures of violence and general risk in comparison to measures of sexual risk, except for the facet 1 score for which effect sizes are comparable for all types of risk assessed. These results suggest that the lifestyle and antisocial behavior content are likely to contribute more to the overall association between PCL measures and risk than the more interpersonal and affective content of the measure. This is interesting as the interpersonal and affective component of the PCL are the aspects that tend to be considered the most different from previously existing measures for risk assessment and are thought to have potential utility in identifying particularly sadistic or violent individuals who are unlikely to be rehabilitated. These findings do, indeed, suggest that this component of the PCL measures is less identified by other risk assessment measures, however, considering that factor 2 and facet 4 scores have been demonstrated to be the strongest predictors of recidivism in the past, these findings may suggest that the components of psychopathy not already accounted for by existent risk assessment measures may not identify risky or violent behavior to the extent previously hypothesized, This supports the suggestion that PCL measures may not add incremental prediction to already existent measures, but rather may be better used as a confirmatory measure for data already gleaned from established risk measures.

Measures of sexual risk followed a similar pattern, with larger effect sizes for total, Factor 2, facet 3, and facet 4 scores and more modest effect sizes for Factor 1, facet 1 and facet 2 scores. However, effect sizes for measures of sexual risk, and the effect size for the association between measures of domestic violence and PCL total score, tended to be more modest overall than those for general and violence risk. It is noteworthy that the difference in effect sizes between sexual risk and facet 1 and facet 2 scores are not significant, as evidenced by overlapping confidence intervals, whereas the difference in effect sizes between the other scores are significantly different. This suggests, perhaps, that sexual risk measures rely similarly heavily on the lifestyle and antisocial behavior components of psychopathy (facet 3 and facet 4) , less heavily on the interpersonal and affective content of psychopathy .

Specific Risk Instruments

Overall, when looking more closely at specific risk assessment measures, effect sizes were remarkably similar and followed the same pattern as the overall effects observed across all studies. Factor 2, facet 3, and facet 4 effect sizes for some specific risk measures approached or exceeded .70 (e.g., r = .78 for PCL Factor 2 scores and SORAG total scores, and r = .64 for PCL facet 3 scores and SAVRY total scores, and r = .72 for PCL facet 4 scores and VRAG total scores). This suggests that the PCL is basically redundant with many, if not all risk assessment measures that are already in existence, especially when you consider these specific sections (lifestyle and antisocial) to PCL measures. Notably, however, effect sizes were more modest for the Static-99 (i.e., moderate effect sizes for PCL total and Factor 2 scores and small effect size for PCL Factor 1 scores). Since the Static-99 measures mostly historical information that related

to risk of sexual reoffending (e.g., number of previous sexual and non-sexual charges and convictions, and demographics of previous victim of sexual crimes) it falls to reason that this measure will substantially overlap with facet 4 content from the PCL measures which also relate to historical information. This measure, however, fails to capture the risk-relevant variance that the personality-oriented, dynamic facets of the PCL confer, and thus will overlap less substantially with other facets, including facet 3 (lifestyle characteristics). Additionally, in contrast with other specific instruments, the effect sizes for the PCL facet 1 and facet 2 scores for the SVR-20 were large suggesting this measure, a sexual violence risk assessment measure, may rely more strongly on interpersonal and affective traits of an offender when considering risk than other measures which rely more strongly on impulsive and antisocial behaviors.

In line with previous research (see for example Churcher et al, 2016 and Stockdale et al, 2014), I expected facet 3 and facet 4 scores, to correlate more strongly with risk assessment measures. Larger effect sizes were expected due to overlap with clinical aspects of risk assessment measures including impulsivity, irresponsibility, and historical aspects of risk assessment measures including early behavioral problems, juvenile delinquency, and criminal versatility. My findings were in line with this prediction, with facet 3 and facet 4 scores boasting stronger correlations with risk assessment measures than facet 1 and facet 2 scores. These findings suggest that the predictive capabilities of PCL measures are likely to be due to the presence of impulsivity and lifestyle aspects which are also included in risk assessment measure rather than the affective and interpersonal aspects of the PCL measure which are the items considered to be more uniquely characteristic of psychopathic offenders. This provides evidence that

the PCL may not add any unique information above and beyond that suggested by other risk measures. This should be taken into consideration when combining information from the PCL with information gathered from established risk assessment measures. Rather than combining information additively, it may be best to consider the overlap in information as confirmatory.

The more modest effect sizes for the associations between PCL measures and sexual violence measure suggest the PCL measures may capture information that is different from that captures by pre-existing measures for sexual violence. Additionally, there are unique risk factors for sexual violence specifically (e.g., victim characteristics – stranger, male, unrelated victims, deviant sexual interests, sexual self-regulation, and history of sexual abuse; see Casey, 2016 and Starzyk & Marshall, 2003), which do not directly overlap with psychopathy. Notably, many sexual violence specific risk factors are static in nature and/or related to characteristics of previous sexual offenses (see Casey, 2016). Given the weak effects noted for the predictive capabilities of the PCL-R for sexual violence in prior research (r = -.12-.14) this data may provide further evidence that PCL measures give inadequate information to assess for risk for future sexual violence, and thus that these measures should not be used in assessments to determine such risk (see for example: Barbaree et al, 2001; Larsen et al, 2020; Sjostedt & Langstrom, 2002).

Additionally, previous research suggests similar predictive abilities for both actuarial and SPJ measures, so it stands to reason that the PCL scores should correlate similarly with both types of measures. The current study suggested similar effect sizes for associations with PCL Factor 2 and facet 3 scores for both actuarial and SPJ measures.

Effect sizes for associations with PCL total, Factor 1, facet 1 and facet 2 scores were larger for structured professional judgment measures than actuarial measures and effect sizes for associations with PCL facet 4 scores were larger for actuarial measures. Despite differences in effect sizes, similar patterns of larger effect sizes for PCL Factor 2, facet 3, and facet 4 scores and smaller effect sizes for PCL Factor 1, facet 1, and facet 2 scores. As noted above, these data suggest that PCL measures may not give unique information rather overlap significantly with the impulsive and lifestyle data already gathered with the use of actuarial and SPJ measures. It also provides evidence that structured professional judgment measure assess for more personality based traits than actuarial measures which focus more highly on assessing for static, historical or lifestyle factors for risk.

Moderator Effects

Analyses indicated a high degree of heterogeneity in effect sizes across studies. However, only a small number of moderators explained a significant amount of variability in the correlation values across studies. Sex, age, and race composition, as well as method of scoring of the PCL measure, type of publication from which data were ascertained, and year of publication of the study were not significant moderators of total score, factor score of facet score meta-analyses.

Analyses considering differences in effect sizes based on the location of the study revealed a significant moderator for correlations between PCL total scores and risk assessment total scores. In these analyses, studies conducted in the European countries and Canada had larger effect sizes than those conducted in the United States. This suggests that the results of my meta-analysis and of the scoring of either the PCL, the risk assessment, or both may be impacted by the country in which these measures are scored.

This finding could be an indication that care should be taken in implementing risk assessment and/or psychopathy measures cross-culturally. Although we do not know the exact reason for this difference, it could be because of different terminology used within the legal system, including different laws and standards for convictions or charges, differs throughout the countries of studies impacting the scoring of risk assessment measures in different countries. This difference may also be due to the difference in presentation of psychopathology among different cultural groups which impacts the scoring of psychopathy or risk assessment measures that require information about mental health.

Additionally, the sample type was determined to be a significant moderator for the correlations between PCL Factor 2 scores and risk measure total scores with larger effect sizes for samples consisting of individuals on probation or parole (r = .74) than those consisting of individuals in correctional settings (r = .66) or institutionalized participants (r - .54). These findings may be a result of range restriction of PCL scores within certain populations of individuals. For example, there is likely to be a larger representation of individuals with psychopathic traits in correctional facilities or on probation than in institutional populations as there are many reasons individuals may be institutionalized rather than psychopathic personality traits (e.g., serious mental illness, competency to stand trial, suicidality). As such, there may be a restricted range of PCL scores for individuals seen within these settings which may impact the results for correlations with risk assessment scores in these populations.

The reliability coefficient, Cronbach's alpha for the PCL scores was determined to be a significant moderator for correlations between PCL Factor 2 scores and risk measure total scores with greater correlations between measures when the internal

consistency of items on Factor 2 was higher. Unreliability constrains the magnitude of possible correlation coefficients (Ponterotto & Ruckdeschel, 2007). Thus, it makes sense that a lower correlation between Factor 2 scores and risk assessment measure total scores would be observed considering the lower reliability coefficients for this component of the PCL.

Implications for Practice

The main questions of this study are designed to help inform practitioners in the field of forensic psychology in the selection of tools utilized to assess for risk of reoffending as well as to inform practitioners on the overlap between measures that may impact how they weight information when making their final risk estimates. Overall, the meta-analytic results showed moderate to large correlations between risk measures and PCL total, Factor 2, facet 3, and facet 4 scores. Correlations between risk measures and PCL Factor 1, facet 1, and facet 2 were consistently smaller. This seems to provide support for the utility of PCL total, Factor 2, facet 3 and facet 4 scores in risk assessment; however, given the substantial effect sizes (in the range of r = .60-.70) it suggest that these PCL scores may not provide unique information about offending patterns, but rather confirm information already gathered by many established risk assessment measures, including information on antisocial, historical features of offending and impulsive and parasitic lifestyle features. The current results also suggest that Factor 1, facet 1, and facet 2 scores may provide some unique information that is not assessed by current risk measures. Although these measures do correlate modestly with risk assessment measures, it is possible that this correlation is related to the overlap with items assessed in Factor 2 scores. The unique associations of Factor 1, facet 1, and facet 2

with risk after accounting for shared variance with Factor 2 should be investigated in future studies. Similarly, the utility of this potentially unique information to provide incremental validity in predicting risk outcomes over and above risk assessments should also be considered in future research, as it was not possible to investigate in the current report. At the same time, it is quite plausible that the PCL will not provide incremental risk prediction. It is important to recall that although this information may be most unique to psychopathic offenders, these scores have in the past been shown to be the least predictive of recidivism (for example, Murrie and colleagues, 2012) and may not be useful in providing information relevant to the assessment of risk of reoffending.

The question is, then, which PCL scores, if any, should we rely on when conducting risk assessments? In the past, Factor 2, facet 3, and facet 4 scores have been shown to be the most predictive of recidivism when considering all possible PCL scores (see for example Chucker et al, 2016; Neal et al, 2015; Stockdale et al, 2014; Warren et al, 2005). Current findings support the link between PCL Factor 2, facet 3, and facet 4 scores in recidivism due to large effect sizes when comparing these scores with scores on established risk assessment measures. That said, it is important to note that because of the high correlation between these measures, it is unlikely that they provide any unique information above and beyond the information gathered from extant risk assessment measures. Considering the high redundancy of these PCL scores and the time required complete the interviews and file review required for the accurate scoring of PCL measures, current evidence suggests that evaluators may more easily gather risk information needed using other established measures. Should the PCL measure be used alongside other measures of risk, it is important to keep in mind the large redundancy

between measures so that evaluators do not fall prey to using the information as additive indication of increased risk, but rather confirmatory of the information gathered by other instruments. It would also be important to consider the potential negative consequences of introducing the label psychopathy into a risk assessment, as it is a highly stigmatizing label that has a strong influence on jurors. Evaluators should carefully weigh the benefit of bringing psychopathy results to an evaluation with the prejudicial effect of the label within the criminal justice system. It is also important to recall, that the current research did not consider incremental validity, so it is premature to definitively suggest that the PCL not be used in risk assessment evaluations. Although current correlations suggest a low possibility for unique contribution, it is still important for research to investigate and find support for this hypothesis.

Limitations

One limitation for the present student was the reliance on only published results for the meta-analytic procedures. It is possible that there are several studies with null finding for the correlations between PCL scores and risk measures scores that could have impacted the results of this study. Results from studies subject to the file-drawer effect could also have impacted the significance of many of the moderators examined in this paper, especially considering some groups had to be excluded from analysis due to small amounts of studies that reported on certain demographic information (e.g., participant race, age, and sex). However, measures of potential publication bias utilized throughout this study consistently suggested that publication bias was not a threat to the results found within the study. Additionally, we did not consider the quality of the studies that were including in this meta-analysis. Had we considered aspects of quality of study such as

population, assessment methodology and training and comparison/analysis, it is possible that studies may have been weighted differently or excluded altogether from the study.

Another limitation of this study is the lack of explanatory moderator variable for the variation observed in many of the meta-analyses. Many of the studies that were included had demographic information that was unreported in at least one domain. Although some of this information was able to be ascertained by contacting authors, some was not due to limitations in the data retrieved or coded for the original study. It is possible that had more information on demographics, study characteristics, and coding practices (e.g., training in the measures administered; information regarding scoring of the PCL based on interview or file review alone) been retrieved these moderators may have been significant. Because few of the moderators analyzed were significantly related to the systematic change in effect sizes between studies, it would be useful to code and analyze additional moderators that may explain some of the differences in findings.

A further limiting factor in this study was the small amount of empirical research reporting correlations between PCL scores and specific risk assessment measures. Although there is currently in excess of 100 different risk assessment measures available for practitioners to decide between when conducting risk assessments, sufficient data was only ascertained to consider the correlations between PCL total and factor scores and risk measure total scores for eight specific risk assessment measures. Even less information was available for the correlations between PCL facet scores and risk measure total scores allowing for meta-analytic findings for only six specific risk measures.

Notwithstanding these limitations, findings from the present study represent an important contribution to the literature on the utility of PCL measures for risk assessment.

The current research provides evidence that although various components of PCL measures, including the PCL total, Factor 2, facet 3, and facet 4 scores, have historically demonstrated moderate predictive abilities as a risk assessment instrument they are also highly redundant with information gathered from already established risk assessment measures. Indeed, those components that appear to contribute the most unique information, the PCL Factor 1, facet 1, and facet 2 scores, have historically demonstrated low predictive abilities as a risk assessment measure. Altogether, the information gathered throughout this study indicate that the PCL should, if used in risk assessments, be considered as a piece of additional evidence to confirm the findings of already existing risk assessment measures rather than as a measure that adds novel information to these measures. Given the time required to administer, review collateral for and score PCL measures, the PCL measures may not be the most ideal to utilize in this manner and their exclusion from risk assessments in which other risk assessment measures with more empirical backing for this purpose should be considered pending additional research.

REFERENCES

Abbiati, M., Palix, J., Gasser, J., & Moulin, V. (2019). Predicting physically violent misconduct in prison: A comparison of four risk assessment instruments. *Behavioral Sciences & the Law, 37*(1), 61-77. http://doi.org/10.1002/bsl.2364

Archer, R. P., Buffington-Vollum, J. K., Stredny, R. V., & Handel, R. W. (2006). A survey of psychological test use patterns among forensic psychologists. *Journal of Personality Assessment, 87*(1), 84-94. http://doi.org/10.1207/s15327752jpa8701_07

Barbaree, H. E., Seto, M. C., Langton, C. M., & Peacock, E. J. (2001). Evaluating the predictive accuracy of six risk assessment instruments for adult sex offenders. *Criminal Justice and Behavior, 28*(4), 490-521. http://doi.org/10.1177/009385480102800406

Bengston, S., & Langstrom, N. (2007). Unguided clinical and actuarial assessment of re-offending risk: A direct comparison with sex offenders in Denmark. *Sexual Abuse, 19*, 135-153. http://doi.org/10.1007/s11194-007-9044-5

Boccaccini, M. T., Chevalier, C. S., Murrie, D. C., & Varela, J. G. (2017). Psychopathy checklist-revised use and reporting practices in sexually violent predator evaluations. *Sexual Abuse: A Journal of Research and Treatment, 29*(6), 592-614. http://doi.org/10.1177/1079063215612443

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* Wiley.

Brazil, K. J., & Forth, A. (2016). Psychopathy Checklist: Youth version (PCL:YV). In: V. Zeigler-Hill & T. Shackelford (Eds.), *Encyclopedia of personality and individual differences* (pp.1-5). Springer.

Campbell, M. A., French, S. & Gendreau, P. (2009). The prediction of violence in adult offenders: A meta-analytic comparison of instruments and methods of assessment. *Criminal Justice and Behavior, 36*(6), 567-590. http://doi.org/10.1177/0093854809333610

Caperton, J. D. (2005). *Predicting recidivism among sex offenders: Utility of the Static-99, Minnesota sex offender screening tool-revised, and psychopathy checklist-revised* (Publication No. 305372844) [Doctoral Dissertation, Sam Houston State University]. ProQuest Dissertations and Theses Global.

Casey, S. (2016). Dynamic risk and sexual offending: The conundrum of assessment. *Psychology, Crime & Law, 22*(1-2), 104-123. http://dx.doi.org/10.1080/1068316X.2015.1111366

Cronbach, L. K., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281-302. https://doi.org/10.1034/h0040957

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243*(4899): 1668-1674. www.jstor.org/stable/1703476.

Desmarais, S. L., Johnson, K. L., & Singh, J. P. (2016). Performance of recidivism risk assessment instruments in U.S. correctional settings. *Psychological Services, 13*(3), 206-222. http://doi.org/0.1037/ser0000075

de Vogel, V., Bruggeman, M., & Lancel, M. (2019). Gender-sensitive violence risk assessment: Predictive validity of six tools in female forensic psychiatric patients.

Criminal Justice and Behavior, 46(4), 528-549. http://doi.org/10.1177/0093854818824135

Douglas, K. S., Yeomans, M., & Boer, D. P. (2005). Comparative validity analysis of multiple measures of violence risk in a sample of criminal offenders. *Criminal Justice and Behavior, 32*(5), 479-510. http://doi.org/10.1177/0093854805278411

Edens, J. F., & Campbell, J. S. (2007). Identifying youths at risk for institutional misconduct: A meta-analytic investigation of the psychopathy checklist measures. *Psychological Services, 4*(1), 13-27. http://doi.org/10.1037/1541-1559.4.1.13

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

Gardner, B. O., Boccaccini, M. T., & Murrie, D. C. (2018). Which PCL-R scores best predict forensic clinicians' opinion of offender risk? *Criminal Justice and Behavior, 45*(9), 1404-1419. http://doi.org10.1177/0093854818789974

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C.. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12*(1), 19-30. http://doi.org/10.1037/1040-3590.12.1.19

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

Hare, R. D. (1998). The Hare PCL-R: Some issues concerning its use and misuse. *Legal and Criminological Psychology, 3*(1), 99-119. http://doi.org/10.1111/j.2044-8333.1998.tb00353.x

Hare, R. D. (1999). Psychopathy as a risk factor for violence. *Psychiatric Quarterly, 70*(3), 181-197. http://doi.org/10.1023/A:1022094925150

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

Hare, R. D., & Neumann, C. S. (2009). Psychopathy: Assessment and forensic implications. *The Canadian Journal of Psychiatry, 54*(12), 791-802.

Harris, D. A., Pedneault, A., & Knight, R. A. (2013). An exploration of burglary in the criminal histories of sex offenders referred for civil commitment. *Psychology, Crime, & Law, 19*(9), 765-781. http://doi.org/10.1080/1068316X.2012.678850

Hart, S. D., Cox, D. N., & Hare, R. D. (1995). *Hare Psychopathy Checklist – Screening version (PCL:SV)*. Multi-Health Systems.

Hawes, S. W., Boccaccini, M. T., & Murrie, D. C. (2013). Psychopathy and the combination of psychopathy and sexual deviance as predictors of sexual recidivism: Meta-analytics findings using the psychopathy checklist-revised. *Psychological Assessment, 25*(1), 233-243. http://doi.org/10.1037/a0030391

Helmus, L., Hanson, R. H., Babchishin, K. M., & Thornton, D. (2015). Sex offender risk assessment with the Risk Matrix 2000: Validation and guidelines for combining with the STABLE-2007. *Journal of Sexual Aggression, 21*(2), 136-157. http://doi.org/10.1080/13552600.2013.870241

Hemphill, J. F., & Hare, R. D. (2004). Some misconceptions about the Hare PCL-R and risk assessment: A reply to Gendreau, Goggin, and Smith. *Criminal Justice and Behavior, 31*(2), 203-243. http://doi.org/10.1177/0093854803261326

Hill, A., Rettenberger, M., Habermann, N., Berner, W., Eher, R., & Briken, P. (2012). The utility of risk assessment instruments for the prediction of recidivism in

sexual homicide perpetrators. *Journal of Interpersonal Violence, 27*(18), 3553-3578. http://doi.org/10.1177/0886260512447570

Hogan, N. R., & Sribney, C. L. (2019). Combinging Static-99R and STABLE-2007 risk categories: An evaluation of the five-level system for risk communication. *Sex Offender Treatment, 14*(1). http://www.sexual-offender-treatment.org/187.html

Jackson, R. L., & Hess, D. T. (2007). Evaluation for civil commitment of sex offenders: A survey of experts. *Sexual Abuse: Journal of Research and Treatment, 19*(4), 425-448. http://doi.org/10.1007/s11194-007-9062-3

JASP Team (2022). JASP (Version 0.16.3) [Computer software].

Kennealy, P. J., Skeem, J. L., Walters, G. D., & Camp, J. (2010) Do core interpersonal and affective traits of PCL-R psychopathy interact with antisocial behavior and disinhibition to predict violence? *Psychological Assessment, 22*(3), 569-580. http://doi.org/10.1037/a0019618

Langevin, R., & Curnoe, S. (2011). Psychopathy, ADHD, and brain dysfunction as predictors of lifetime recidivism among sex offenders. *International Journal of Offender Therapy and Comparative Criminology, 55*(1), 5-26. http://doi.org/10.1177/0306624X09360968

Langton, C. M. (2003). *Contrasting approaches to risk assessment with adult male sexual offenders: An evaluation of recidivism prediction schemes and the utility of supplementary clinical information for enhancing predictive accuracy* [Doctoral dissertation, University of Toronto]. ProQuest Dissertations and Theses Global.

Larsen, R. R., Jalava, J. & Griffiths, S. (2020) Are psychopathy checklist (PCL) psychopaths dangerous, untreatable, and without conscience? A systematic review

of the empirical evidence. *Psychology, Public Policy, and Law, 26*(3), 297-311. http://doi.org/10.1037/law0000239

Leistico, A. R., Salekin, R. T., DeCoster, J., & Rogers, R. (2008). A large-scale meta-analysis relating the Hare measures of psychopathy to antisocial conduct. *Law and Human Behavior, 32*(1), 28-45. http://doi.org/10.1007/s10979-007-9096-6

Melton, G. B., Petrila, J., Poythress, N. G., Slobgin, C., Otto, R. K., Mossman, D., & Condie, L. O. (2018). *Psychological evaluations for the courts: A handbook for mental health professional and lawyers* (4th ed., pp. 260-312). The Guilford Press.

Monahan, J. (1981). The clinical prediction of violent behavior. *Crime & Delinquency Issues: A Monograph Series, ADM 81-921*, 134.

Mokros, A., Vohs, K., & Habermeyer, E. (2014). Psychopathy and violent reoffending in German-speaking countries. *European Journal of Psychological Assessment, 30*(2), 117-129. http://doi.org/10.1027/1015-5759/a000178

Murrie, D. C., Boccaccini, M. T., Caperton, J., & Rufino, K. (2012). Brief report: Field validity of the psychopathy checklist-revised in sex offender risk assessment. *Psychological Assessment, 24*(2), 524-529. http://doi.org/10.1037/a0026015

Neal, T. M. S., & Grisso, T. (2014). Assessment practices and expert judgment methods in forensic psychology and psychiatry: An international snapshot. *Criminal Justice and Behavior, 41*(12), 1406-1421. http://doi.org/10.1177/0093854814548449

Neal, T. M. S., Miller, S. L., & Shealy, R. C. (2015). A field study of a comprehensive violence risk assessment battery. *Criminal Justice and Behavior, 42*(9), 952-968. http://doi.org/10.1177/0093854815572252

Parent, G., Guay, J., & Knight, R. A. (2011). An assessment of long-term risk of recidivism by adult sex offenders: One size doesn't fit all. *Criminal Justice and Behavior, 38*(2), 188-209. http://doi.org/10.1177/0093854810388238

Ponterotto, J. G., & Ruckdeschel, D. E. (2007). An overview of coefficient alpha and a reliability matrix for estimating adequacy of internal consistency coefficients with psychological research measures. *Perceptual and Motor Skills, 105*(3), 997-1014. http://doi.org/10.2466/PMS.105.3.997-1014

Povli, N. H. (1999). *The prediction of violence in pre-trial forensic patients: The relative efficacy of statistical versus clinical predictions of dangerousness*. [Doctoral dissertation, Simon Fraser University]. ProQuest Dissertations and Theses Global.

Scalora, M. J., Viñas-Racionero, R., & Cawood, J. S. (2020). Comparison of the HCR-20V3, the WAVR-21V3, and the CAG performance across workplace homicide scenarios: A pilot study. *Journal of Threat Assessment and Management, 7*(3-4), 186-199. http://doi.org/10.1037/tam0000154

Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review, 31*(3), 499-513. http://doi.org/10.1016/j.cpr.2010.11.009

Singh, J. P., Desmarais, S. L., Hurducas, C., Arbach-Lucioni, K., Condemarin, C., Dean, K., Doyle, M., Folino, J. O., Godoy-Cervera, V., Grann, M., Ho, R. M. Y., Large, M. M., Nielsen, L. H., Pham, T. H., Rebocho, M. F., Reeves, K. A., Rettenberger, M., de Ruiter, C., Seewald, K., & Otto, R. K. (2014). International perspectives on the practical application of violence risk assessment: A global survey of 44

countries. *International Journal of Forensic Mental Health, 13*(3), 193-206. http://doi.org/10.1080/14999013.2014.922141

Sjostedt, G., & Langstrom, N. (2002). Assessment of risk for criminal recidivism among rapists: A comparison of four different measures. *Psychology, Crime and Law, 8*(1), 25-40. http://doi.org/10.1080/10683160208401807

Smid, W. J., Kamphuis, J. H., Wever, E. C., & Van Beek, D. J. (2014) A comparison of the predictive properties of nine sex offender risk assessment instruments. *Psychological Assessment, 26*(3), 691-703. http://doi.org/10.1037/a0036616

Starzyk, K. B., & Marshall, W. L. (2003). Childhood family and personological risk factors for sexual offending. *Aggression and Violent Behavior, 8*(1), 93-105. http://dx.doi.org/10.1016/S1359-1789(01)00053-2

Sturup, J., Karlberg, D., Fredriksson, B., Lihoff, T., & Kristiansson, M. (2016). Risk assessments and recidivism among a population-based group of Swedish offenders sentenced to life in prison. *Criminal Behaviour and Mental Health, 26*(2), 124-135. http://doi.org/10.1002/cbm.1941

Texas Health & Safety Code §841.022 (2015) https://statutes.capitol.texas.gov/Docs/HS/htm/HS.841.htm

Viljoen, J. L., Elkovitch, N., Scalora, M. J., & Ullman, D. (2009). Assessment of reoffense risk in adolescents who have committed sexual offenses: Predictive validity of the ERASOR, PCL:YV, YLS/CMI, and Static-99. *Criminal Justice and Behavior, 36*(10), 981-1000. http://doi.org/10.1177/0093854809340991

Viljoen, J. L., McLachlan, K., & Vincent, G. M. (2010). Assessing violence risk and psychopathy in juvenile and adult offenders: A survey of clinical practices. *Assessment, 171*(3), 377-395. http://doi.org/10.1177/1073191109359587

Walters, G. D. (2003). Predicting criminal justice outcomes with the psychopathy checklist and lifestyle criminality screening form: A meta-analytic comparison. *Behavioral Sciences and the Law, 21*(1), 89-102. http://doi.org/10.1002/bsl.519

Welsh, J. L., Schmidt, F., McKinnon, L., Chattha, H. K., & Meyers, J. R. (2008). A comparative study of adolescent risk assessment instruments: Predictive and incremental validity. *Assessment, 15*(1), 104-115. http://doi.org/10.1177/1073191107307966

Yang, M., Wong, S. C. P., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin, 136*(5), 740-767. http://doi.org/10.1037/a0020473

VITA

Stephanie L. Goodwin, B.S

CURRICULUM VITAE
Sam Houston State University | Huntsville, TX

EDUCATION

Doctor of Philosophy Student | Clinical Psychology (Forensic Emphasis) In Progress
Sam Houston State University - Huntsville, TX (APA Accredited)
Advisor: Marcus Boccaccini, Ph.D.
Research Focus: Forensic Psychology & Risk Assessment

Bachelor of Science | Psychology (Summa Cum Laude) 2017
Saint Mary's University – Halifax, NS
Advisor: Marguerite Ternes, Ph.D
Thesis: *Verbal Credibility in Memory of Emotionally Negative Events*

Diploma | Forensic Sciences 2017
Saint Mary's University – Halifax, NS

Bachelor of Science | Biochemistry and Molecular Biology 2014
Dalhousie University – Halifax, NS

WORK EXPERIENCE

Graduate Assistant/Research Assistant 2019-Present
Sam Houston State University – Huntsville, TX
Setting: Higher Education
Population: University students and community members
Supervisor: Marcus Boccaccini, Ph.D.
Description: Conduct research on risk assessment and forensic psychology related topics and provide instruction in introductory psychology to undergraduate students.

Psychology Technician 2018-2019
Forensic Sexual Behavior Program, Nova Scotia Health Authority – Halifax, NS
Setting: Community Mental Health Center
Population: Individuals who had been convicted of a sexual offense, or community members looking to undergo a sex offender risk assessment.
Supervisor: Angela Connors, Ph.D.
Description: Penile plethysmography assessments for male sex offenders or those undergoing a community forensic assessment, scoring MMPI-2, MCMI, PDS, STAXI, and BDI-2 assessments, maintenance of an assessment and treatment database, and coding of recidivism data for past program clients for research purposes.

Research Assistant 2017-2018
Halifax Regional Police – Halifax, NS
Setting: Police Department
Population: Currently employed police officers
Supervisor: Christopher Giacomantonio, Ph.D.
Description: Research related to the effectiveness of verbal de-escalation training for current police officers, assistance in writing a technical review of this research project, and assistance in the development of a tool for assessing the presence of implicit bias pre- and post-training in Fair and Impartial Policing skills.

CLINICAL/PRACTICA EXPERIENCE

Practicum Intern 2021-2022
Center for Clinical and Forensic Psychology – College Station, TX
Setting: Private Practice
Population: Primarily low-income, ethnically diverse adults
Supervisor: Jennifer Rockett, Ph.D, LP
Description: Complete a variety of forensic assessments, including competency to stand trial assessments, competency to proceed pro se assessments, sanity assessment risk assessments, mitigation assessments, sexually violent predator evaluations, and guardianship evaluations. Opportunities to complete assessments in a variety of different jurisdictions, including Texas, New Mexico, PSYPACT states, and federally. Read collateral data and complete collateral interviewing calls with individuals familiar with the defendant. Write reports following assessments and provide professional opinions to the courts and/or attorneys.

Student Clinician/Co-Therapist 2022-Present
TEAM Forensic Services – Huntsville, TX
Setting: Private Practice
Population: Primarily low-income, rural, ethnically diverse adult males on probation or parole for sexual offenses
*Supervisor:*Holly Miller, Ph.D, LSOTP
Description: Co-facilitate mandated, evidence-based manualized group treatment with a Licensed Sex Offender Treatment Provider/Licensed Psychologist. Participate in a peer supervisory process group with fellow student clinicians and a licensed psychologist

Clinic Coordinator 2021-2022
Psychological Services Center, Sam Houston State University – Huntsville, TX
Setting: Community Mental Health Center
Population: Ethnically diverse, low-income adolescents, adult college students, and community members with a variety of mental health concerns
Supervisors: Jaime Anderson, Ph.D, Jared Ruchensky, Ph.D., Jorge Varela, Ph.D., Chelsea Ratcliff, Ph.D., & Mary Alice Conroy, Ph.D. ABPP
Description: Conduct phone intakes and manage clients for student clinicians at the PSC. Conduct Quality assurance checks on a semester basis and close files of completed assessment and therapy clients. Lead clinic meetings and coordinate with the clinic director and administrative staff about information to be communicated to student

clinicians. Conduct community-based therapy and assessment services for children, adolescents, and adults. Collaborate on treatment planning and closely monitor treatment goals and progress.

Student Clinician 2020-2021
Psychological Services Center, Sam Houston State University – Huntsville, TX
Setting: Community Mental Health Center
Population: Ethnically diverse, low-income adolescents, adult college students and community members with a variety of mental health concerns
Supervisors: Craig Henderson, Ph.D., Laura Drislane, Ph.D., Mary Alice Conroy, Ph.D. ABPP
Description: Conduct community-based therapy and assessment services for children, adolescents, and adults. Collaborate on treatment planning and closely monitor treatment goals and progress.

SUPERVISORY EXPERIENCE

Peer Supervisor 2021-2022
Psychological Services Center, Sam Houston State University – Huntsville, TX
Beginning Doctoral Practicum, Doctoral Practicum I
Supervisees: First-year and Second-year doctoral student clinicians
Supervisor: Mary Alice Conroy, Ph.D. ABPP
Description: Co-facilitate supervision sessions with a licensed psychologist, review mock sessions with supervisee, and provide feedback related to interviewing techniques and therapy skills.

TEACHING EXPERIENCE

Lecturer Fall 2022
Sam Houston State University – Huntsville, TX
Course: PSYC 3383 Psychology and the Law
Description: Created syllabus and lecturers for a fall term course and lectured using an asynchronous online model. Created, disseminated, and graded assignments, discussion questions, and examinations. Gained experience using Blackboard to input grades and upload course presentations, and Kaltura to record lectures.

Lecturer Fall 2020-Spring 2021
Sam Houston State University – Huntsville, TX
Course: PSYC 1301 Introduction to Psychology
Description: Created syllabus and lecturers for a fall and spring term course and lectured using a hybrid (in-person and zoom) model. Created, disseminated, and graded assignments and examinations. Gained experience using Blackboard to input grades and upload course presentations.

RESEARCH EXPERIENCE

Graduate Research Assistant 2019-Present
Sam Houston State University – Huntsville, TX
Department of Psychology and Philosophy

Supervisor: Marcus Boccaccini, Ph.D.
Description: Conduct various research projects, create poster presentations and write manuscripts for completed projects.
Current Projects: Identifying Influential Risk Score and Diagnosis Combinations in SVP Evaluations (Co-Investigator), The Association Between Structured Professional Judgment Measure Total Scores and Summary Risk Ratings: Implications for Predictive Validity (Co-Investigator).

Undergraduate Research Assistant 2015-2017
Saint Mary's University – Halifax, NS
Department of Psychology
Supervisor: Marguerite Ternes, Ph.D.
Description: Conduct various research projects, create poster presentations and write manuscripts for completed projects.
Projects: Two Truths and a Lie (Research Assistant), Verbal Credibility in Memory of Emotionally Negative Events (Principal Investigator; Honors Thesis).

Undergraduate Research Assistant 2011
Dalhousie University – Halifax, NS
Department of Psychology
Supervisors: Timothy O'Leary & Richard Brown, Ph.D.
Description: Conduct a student mini-project involving the coding of cues of locomotor deficits and anxiety in mouse models of Alzheimer's disease and develop a post presentation of the findings.

PUBLICATIONS

Giacomantonio, C., **Goodwin, S.** (2019). Learning to de-escalate: Evaluating the behavioural impact of Verbal Judo training on police constables. Submitted to Police Practice and Research: An International Journal.

Ternes, M., **Goodwin, S.**, Hyland, K. (2018). Substance Use Disorders in Correctional Populations, in *The Practice of Correctional Psychology.*

CONFERENCE PRESENTATIONS

Goodwin, S., Boccaccini, M., & Harris., P. (2020). Identifying Influential Risk Score and Diagnosis Combinations in SVP Evaluations. Poster presented at the Annual Meeting of the American Psychology-Law Society, New Orleans, LA.

Goodwin, S., Ternes, M. (2017). Verbal credibility in memory of emotionally negative events. Paper presented at the Forensic Psychology in Canada Conference, Carleton University, Ottawa, ON.

Goodwin, S., Ternes, M. (2017). Verbal credibility in memory of emotionally negative events. Paper presented at the Saint Mary's University Psychology Student Conference, Halifax, NS.

Assal, A., Beadle, T., Campbell, M., Choi, A., **Goodwin, S.**, & Rumbolt, K. (2011). Investigating anxiety and locomotor activity in the 5XFAD mouse model of Alzheimer's disease. Poster presented at the Dalhousie Poster Presentation, Dalhousie Integrated Science Program, Halifax, NS.

PROFESSIONAL SERVICE AND LEADERSHIP POSITIONS

Executive Officer: Treasurer 2022-Present
Graduate Student Psychology Organization
Sam Houston State University

Member at Large, Member of Mentorship Subcommittee 2022-Present
Clinical Psychology Doctoral Program Diversity Committee
Sam Houston State University

Executive Officer: President 2021-2022
Graduate Student Psychology Organization
Sam Houston State University

Executive Officer: Secretary 2021-2022
Clinical Psychology Doctoral Program Diversity Committee
Sam Houston State University

Student Representative 2020-2021
Clinical Psychology Doctoral Program
Sam Houston State University

Executive Member: Events Coordinator 2016-2017
Forensic Science Society
Saint Mary's University

COMMUNITY SERVICE INVOLVEMENT

CASA Volunteer/Guardian Ad Litem 2021-Present
CASA of Walker, Trinity, and San Jacinto Counties – Huntsville, TX
Description: Visit child(ren) involved in CPS cases and check in on their current status and needs. Contact various spheres of the child(ren)'s life including CPS, parents, placements, medical and educational professionals regarding progress and needs. Compile information into court reports that speak to the child(ren)'s best interests, attend court hearings, participate in family conferences, and participate in collaborative family engagement sessions.

PROFESSIONAL DEVELOPMENT AND TRAINING

Haven 101 2020
Sam Houston State University

CORE Certificate 2020

CITI Program

2020 Physicians for Human Rights Asylum Network Training at BCM 2020
Baylor College of Medicine – Houston, TX

Cardiopulmonary Resuscitation (CPR) and First Aid 2018
Nova Scotia Health Authority

Psychopathy Checklist-Revised (PCL-R) 2017
Forensic Psychology in Canada Conference – Ottawa, ON
Adele Forth, Ph.D.

Criteria-Based Content-Analysis (CBCA) 2016
Saint Mary's University – Halifax, NS
Marguerite Ternes, Ph.D.

Step-Wise Interview Training 2016
Saint Mary's University – Halifax, NS
Marguerite Ternes, Ph.D.

Fire Forensics 2015
Nova Scotia Firefighters School – Halifax, NS
Greg Olson, M.S.

Forensic Photography 2014
Saint Mary's University – Halifax, NS
Michael Creagen & Lisa Pessin

PROFESSIONAL MEMBERSHIPS

Student Member, American Psychological Science (APS) 2019-Present
Student Member, American Psychological Association (APA) 2019-Present
Student Member, American Psychology-Law Society (AP-LS) 2016-Present
Student Member, Canadian Psychological Association (CPA) 2016-2019

HONORS, AWARDS, AND SCHOLARSHIPS

Good Neighbor Scholarship 2022-2023
Office of International Programs, Sam Houston State University

Good Neighbor Scholarship 2020-2021
Office of International Programs, Sam Houston State University

Graduate Student Scholarship 2019-Present
Sam Houston State University

Undergraduate Academic Achievement Scholarship 2015-2016
Saint Mary's University

Father JJ Hennessey Memorial Scholarship 2016
Saint Mary's University

Dean's List 2014-2017
Saint Mary's University

J&W Murphy Scholarship 2010
Dalhousie University

Entrance Scholarship 2010
Dalhousie University

LANGUAGES

- English (fluent)
- French (proficient)