CLINICAL UTILITY OF PSYCHOPATHY SUBTYPES BASED ON LATENT PROFILE ANALYSIS

A Dissertation

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

Doctor of Philosophy

---

by

Gabriele Trupp

August, 2022

CLINICAL UTILITY OF PSYCHOPATHY SUBTYPES BASED ON LATENT PROFILE ANALYSIS

by

Gabriele Trupp

____

APPROVED:

Marcus Boccaccini, PhD
Committee Director

Laura Drislane, PhD
Committee Co-Director

Daniel Murrie, PhD
Committee Member

Craig Henderson, PhD
Committee Member

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

ABSTRACT

Trupp, Gabriele. *Clinical utility of psychopathy subtypes based on latent profile analysis.* Doctor of Philosophy (Clinical Psychology) August 2020, Sam Houston State University, Huntsville, Texas.

Studies using latent profile analysis of the Psychopathy Checklist – Revised (PCL-R) report that groups differ on external correlates relevant to forensic evaluations. Some of these external correlates are being used to support the use of the PCL-R in high stakes forensic settings, such as capital sentencing (Olver et al., 2020). For subtype findings to be useful in practice, clinicians need to be able to reliably assign offenders to correct subgroups.

The current study contained two parts, both of which aimed to address whether individuals can accurately classify PCL-R profiles into their correct subtype as found by subtyping research performed by McCallum and colleagues (2020). In the first part of the study, psychology doctoral students (N = 12) were able to classify profiles with moderate to high accuracy, with some differences based on subtype and whether the data were presented as mean item scores or summed facet scores. The overall difference in accuracy between the mean item and summed facet scores was not statistically significant.

The second part of the study asked clinicians to classify PCL-R profiles into their empirically based subtype. The profiles were presented as a score sheet similar to the Scoring Grid presented in the manual (Hare, 2003) to increase generalizability. Clinicians (N = 37) were better able to classify the prototypic subtype, with more difficultly on the sociopathic and callous-conning subtypes. As prototypicality of the profiles lowered, clinicians had more difficulty with accurate classification.

Overall, this study shows that individuals are better at classifying some subtypes compared to others; however, more research is needed to investigate what differences are causing discrepant classification accuracy between the subtypes. Further, although the overall difference between the two presentation methods in the first study (mean item scores and summed facet scores) was not statistically significant, participants stated that the profiles presented as mean item scores were easier to classify. This may be because they better map on to the way the subtypes are presented in the research literature. Researchers may want to consider presenting their findings in a manner more consistent to what practitioners have in the field to facilitate more effective use of their findings.

KEY WORDS: Psychopathy; Latent profile analysis; Subtyping; Psychopathy Checklist – Revised.

ACKNOWLEDGEMENTS

Thank you to my Co-Chairs. Dr. Boccaccini, I have learned so much from working with you throughout my time in the doctoral program and I hope to continue to use the skills and knowledge you have instilled in me throughout my research career. Dr. Drislane, your knowledge of psychopathy was integral to conducting this study and it was a pleasure working with you. I would also like to thank my family, including Miss J, for their continued love and support. You are my biggest champions and I appreciate everything you have done to help me get to this point.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure **Page**

CHAPTER I

Introduction

Although psychopathy is one of the most frequently-studied topics in psychology and Hare's Psychopathy Checklist-Revised (PCL-R; Hare, 2003) is used extensively by forensic practitioners in the field (see Boccaccini et al., 2017; Neal & Grisso, 2014; Viljoen et al., 2010), there are still large gaps in the application of PCL-R research findings to clinical practice. One of these gaps is at the forefront of an ongoing debate about the use of the PCL-R in capital case assessments (DeMatteo et al., 2020; Olver et al., 2020). A key issue in this debate is the extent to which findings from latent variable models (e.g., structural equation models, latent profile analyses) are useful in the field for making decisions about individual evaluees. Those arguing against the use of the PCL-R in capital cases point to poor field reliability and relatively low predictive validity for PCL-R scores (see DeMatteo et al., 2020), while those arguing in favor of the PCL-R point to findings from latent variable models that show strong performance of the PCL-R after measurement error has been statistically removed from the observed scores (see Olver et al., 2020).

As just one example, PCL-R proponents support their position by citing a study that used latent profile analysis (LPA) with PCL-R facet scores to sort offenders into three subtype groups and found that the three groups differed in their rates of future institutional violence (Olver et al., 2020). Specifically, offenders within the prototypic group (elevations on all four PCL-R facets) and externalizing group (elevations on Factor 2) had significantly more disciplinary reports against persons than individuals in the general offender group (low scores on all facets). Further, offenders within the prototypic group had significantly more security violation disciplinary reports than offenders in the general offender group. Although these types of findings based on models that account for measurement error are useful for studying psychopathy, the extent to which a practitioner can use them to help make scientifically-informed decisions about patients or offenders is unclear. In clinical practice, evaluators do not have access to scores with measurement error removed. Instead, practitioners have field scores (i.e., *observed* scores on PCL-R facets and factors) which always include measurement error. For LPA subtype findings to be useful for clinical practice, clinicians scoring the PCL-R would need to first be able to reliably assign offenders to the correct subtype groups before they could make appropriate generalizations about relevant outcome variables (e.g., future violence, treatment amenability, psychopathology; Klein et al., 2018; Krstic et al., 2018; McCallum et al., 2020).

None of the many existing PCL-R subtype studies have examined whether anyone (e.g., researchers, clinicians, research volunteers) can look at an individual's PCL-R scores and classify them into the correct latent class. If practitioners cannot make these subgroup assignments with a reasonable degree of accuracy, LPA results available in the current published research literature cannot be directly useful for real-world decision-making. My proposed dissertation study attempts to address this gap in the literature by asking both graduate students trained in forensic assessment and licensed forensic practitioners with experience using the PCL-R to classify PCL-R profiles into their LPA-identified subtype groups.

Psychopathy and the Psychopathy Checklist-Revised

The idea of psychopathy has been recognized since the 19th century (Pinel, 1806). Prominent authors such as Cleckley (1941), Karpman (1946, 1948), Hare (1991, 2003), and Patrick (Patrick et al., 2009) have worked toward a more comprehensive understanding of the disorder and there are now at least 12 assessment tools designed specifically to measure psychopathic traits (see Patrick, 2018). However, no one conceptualization has gained universal support among psychopathy scholars and there is still much debate about the etiology of psychopathy and the combination of traits that are central to the construct (see Patrick, 2018).

Despite this lack of consensus in theory and measurement, only the Hare family of measures is used routinely in clinical-forensic practice: Hare's Psychopathy Checklist-Revised (2003), the Psychopathy Checklist: Screening Version (PCL:SV; Hart et al., 1995), and the Psychopathy Checklist: Youth Version (PCL:YV; Forth et al., 2003; see Boccaccini et al., 2017; Neal & Grisso, 2014; Viljoen et al., 2010). To use the PCL-R, clinicians use information from an interview and file-review to score individuals on 20-items that assess affective, interpersonal, lifestyle, and behavioral traits associated with psychopathy. Scores on these items are summed to provide an overall or total psychopathy score. Hare (1991) originally proposed that the PCL-R contained two factors, with Factor 1 consisting of eight items that encompass the interpersonal and affective features of psychopathy and Factor 2 consisting of nine items that represent the more antisocial and socially deviant elements of psychopathy. More recent research suggests that four facets underpin these two factors (Hare, 2003; Hare & Neumann, 2008). Specifically, Facet 1 (Interpersonal) and Facet 2 (Affective) are subsumed by

Factor 1, whereas Facet 3 (Lifestyle) and Facet 4 (Antisocial) are subsumed by Factor 2. There are also two items that contribute to the total score but are not considered a part of any of the factors or facets (Promiscuous Sexual Behavior, Many Short-Term Marital Relationships).

Clinicians use the PCL measures in forensic contexts due to the small- to moderate-sized associations between scores on these measures and clinically meaningful outcome variables. For example, meta-analytic studies suggest that PCL scores, specifically Factor 2 scores, have a moderate association with antisocial conduct (Leistico et al., 2008), institutional misconduct (Guy et al., 2005), general and violent recidivism (Walters, 2003), and reactive violence (Blais et al., 2014). Research also shows Facet 3 scores to be important in explaining both instrumental and reactive aggression (Blais et al., 2014). The PCL-R is often used in sex offender risk assessment (Boccaccini et al., 2017), where Factor 2 and Facet 4 scores are small to moderate predictors of violent and sexual recidivism, and there is some evidence that offenders with high levels of both psychopathy and sexual deviance are at an especially high risk for reoffending (Hawes et al., 2013). These findings are interesting considering the continual debate regarding whether Factor 2 is central to the construct of psychopathy (Hare & Neumann, 2010) or whether the antisocial behavior represented by Facet 4 is simply a manifestation of other more central elements of psychopathy (Skeem & Cook, 2010).

Studies showing that effects for a particular PCL-R factor or facet score are stronger than those for other factors and facets in an individual study or a meta-analysis can be useful when practitioners want to identify the single strongest predictive score from the measure and use only that score. However, interpreting just one facet score from the measure may leave out useful information provided by the other facet scores or by particular configurations of facet scores. For example, those evaluating sex offenders with the PCL-R may miss potentially meaningful information about treatment responsiveness if they focus on only Facet 4 scores. Indeed, researchers have found that Factor 1 scores were more strongly correlated with poor treatment responsivity among sex offenders (e.g., poor insight, attitudes/cognitions supportive of offending, noncompliance with treatment) than Factor 2 scores (Olver & Wong, 2009). In a subsequent study, researchers specifically found Facet 2, the Affective facet, to be the strongest predictor of decreased therapeutic progress and Facet 3, the Lifestyle facet, to be the strongest predictor of treatment noncompliance (Sewall & Olver, 2019). Therefore, a practitioner who decided to focus on only one facet score when making decisions about treatment planning would have overlooked other potentially relevant information from the individual's score profile.

Similarly, interpreting just the PCL-R total score may also lead practitioners to miss potentially meaningful information provided by other scores. Two individuals who have similar total scores can have very different trait profiles at the item, facet, and factor levels. Studies that focus on reporting separate effects for each facet score and meta-analyses that average facet- and factor-level effects across these studies imply that finding and using the single most predictive score is the preferred interpretation approach. There are, however, other types of analyses that allow for a more nuanced interpretation of an individual's entire PCL profile, as opposed to a score-by-score or facet-by-facet approach.

Subtyping and Psychopathy

Subtyping analyses allow psychopathy researchers to consider how individuals score across a group of measures (e.g., facet scores) and to classify them into subgroups with similar score patterns (Hicks & Drislane, 2018). Thus, subtyping classifications provide a means for summarizing distinct configurations of scores across multiple facets simultaneously. They can then compare the subgroups on external correlates to provide information about clinically meaningful differences between the subtypes. Early psychopathy subtyping literature used various forms of cluster analysis, including k-means, Ward's method, and model-based cluster analysis. As subtyping literature expanded, researchers also began using LPA, which seeks to identify discrete, homogeneous subgroups based on the similarity of mean-levels on a set of continuous variables (Williams & Kibowski, 2016). In other words, LPA attempts to group individuals into subtypes based on their responses to a particular set of questions.

There are now more than 26 subtyping studies in the psychopathy literature, with many focusing on identifying the differentiating characteristics of primary and secondary psychopathy using different psychopathy measures (see Hicks & Drislane, 2018). Across studies utilizing diverse samples, as well as different psychopathy assessment tools and data-analytic techniques, research suggests that there are primary and secondary psychopathy subtypes among those with high scores on the measures (e.g., PCL-R $\geq$ 25). Although individuals in both groups obtain relatively high scores on the PCL-R, when compared to one another, the primary subtype has higher PCL-R total and Factor 1 scores, whereas the secondary group has higher Factor 2 scores (Hicks & Drislane, 2018).

Subtyping and the PCL-R

Although early subtyping literature using the PCL-R found support for primary and secondary psychopathy subtypes (Hicks et al., 2004; Olver et al., 2015; Skeem et al., 2007; Swogger & Kosson, 2007; Vassileva et al., 2005), a distinction can be made between studies limiting their sample to participants with high PCL-R scores and studies analyzing scores from all participants. Studies using cluster analysis of just those who score high on psychopathy tend to only find a primary psychopathy cluster and a secondary psychopathy cluster (Hicks et al., 2004; Olver et al., 2015; Skeem et al., 2007). However, studies using cluster analysis on a full range of PCL-R scores find more than two groups, which differ in their severity, ranging from groups that are non-psychopathic to psychopathic, with some groups designated by more moderate psychopathic features (Swogger & Kosson, 2007; Vassileva et al., 2005).

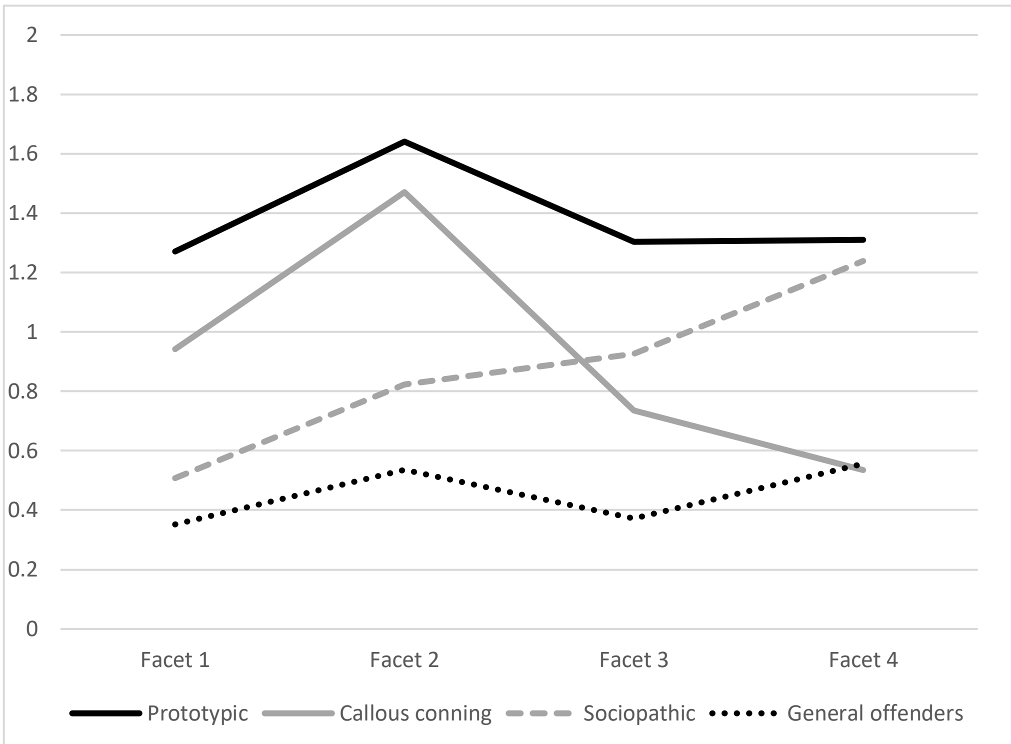
More recently, researchers have moved toward using LPA to investigate PCL-R subtypes. In one study using a subsample of high scoring sex offenders (PCL-R $\geq$ 25), McCallum and colleagues (2020) found support for primary and secondary subtypes. However, in another study also using higher scoring participants, Mokros and colleagues (2015) found evidence for three subtypes that they labeled manipulative, aggressive, and sociopathic. They further explain that the manipulative and aggressive groups are variants of primary psychopathy, whereas the third group is reminiscent of secondary psychopathy. However, as in prior literature using other subtyping analyses, LPA studies using samples of individuals who obtained a range of scores on the PCL-R, not just high scores, suggest a different set of subtypes.

Another approach for identifying subtypes with PCL-R scores has been to use LPA with facet scores from all patients or offenders in a sample, as opposed only those with high psychopathy scores (Neumann et al., 2016). There are now at least seven of these studies. Most of these studies have concluded that there are four PCL-R subtypes: prototypic, callous-conning, sociopathic, and non-psychopathic/general offenders (Hare et al., 2018; Klein Haneveld et al., 2018; Krstic et al., 2018; Lehmann et al., 2019; McCallum et al., 2020; Neumann et al., 2016).

Figure 1 provides a plot for each of the mean item scores on each facet for the four subgroups from a recent LPA study (McCallum et al., 2020). The prototypic subtype is characterized by high average scores across all four PCL-R facets. The callous-conning subtype is characterized by elevated interpersonal and affective traits with comparatively lower lifestyle and antisocial traits. The sociopathic subtype is characterized by elevated lifestyle and antisocial traits with comparatively lower interpersonal and affective traits. Lastly, the non-psychopathic subtype is characterized by offenders who exhibit low to average psychopathic traits across all four facets. Notably, when participants representing a range of scores are included in the analyses, subtypes of primary and secondary psychopathy are not identified by the analysis. Indeed, although LPA models using the full range of scores identify two groups with moderate facet-score elevations similar to primary and secondary psychopathy in some respects (callous-conning and sociopathic, respectively), individuals in these groups do not score high enough on both PCL-R factors for primary and secondary classifications to be wholly accurate (Neumann et al., 2016).

Figure 1

The Four LPA Subtype Profiles Found by McCallum et al. (2020) by PCL-R Mean Item Score on each Facet



Note. Class 1 = prototypic psychopaths (n = 239), Class 2 = callous-conning offenders (n = 154); Class 3 = sociopathic offenders (n = 96), Class 4 = non-psychopathic general offenders (n = 126).

External Correlates

To get a better sense of what these subtypes might mean clinically, researchers have analyzed how those in the subtype groups differ on relevant external correlates. In studies using samples of high psychopathic individuals, researchers have found that primary psychopathy is associated with risk-taking, strategic action, low stress, lack of close attachments, high social dominance (Hicks et al., 2004), comparably better clinical and interpersonal functioning (Skeem et al., 2007), and comparatively lower rates of sexual violence (Olver et al., 2015). In contrast, secondary psychopathy is associated with aggression, high negative emotionality, low constraint (Hicks et al., 2004), more features of comorbid mental disorders (Skeem et al., 2007), and relatively greater criminogenic needs (Olver et al., 2015). Studies using a full distribution of psychopathy scores extend this research with findings that suggest subgroups high on Factor 1 and lower on Factor 2 are characterized by comparably lower anxiety and higher number of violent crimes, whereas subgroups higher on Factor 2 and lower on Factor 1 are characterized by comparably higher drug and alcohol problems (Swogger & Kosson, 2007; Vassileva et al., 2005).

Researchers who have used all psychopathy scores and uncovered four subtypes have found that prototypic offenders differ from the other subtypes in that they score higher on violence risk measures (Lehmann et al., 2019; McCallum et al., 2020; Neumann et al., 2016), display more features of personality disorders (Klein Haneveld et al., 2018; McCallum et al., 2020), have committed more violent offenses (Kristic et al., 2018; Lehmann et al., 2018), and drop out of treatment at a significantly higher rate (Klein Haneveld et al., 2018). Those in the sociopathic subtype tend to show the second highest levels of risk and violence (Krstic et al., 2018; Lehmann et al., 2019; McCallum et al., 2020). Further, among a three-class solution, Neumann and Baskin-Sommers (as cited in Olver et al., 2020) also found prototypic offenders to have more institutional disciplinary reports.

In a recent study (McCallum et al., 2020), researchers found that the callous-conning subgroup and the sociopathic subgroup exhibited very similar mean PCL-R total scores (*M* = 18.12 and 18.22, respectively). However, these groups differed significantly on their responses to the Personality Assessment Inventory (PAI; Morey, 1991).

Specifically, individuals in the callous-conning subgroup reported significantly lower externalizing psychopathology and antisocial features when compared to the sociopathic group. Further, the callous-conning subgroup exhibited lower violence potential scores, problematic treatment scores, and overall levels of impairment and distress. This study emphasizes the importance of analyses that allow researchers and clinicians to examine variants of PCL-R profiles – if a clinician interacting with an individual from this sample had made treatment or legal recommendations based on the PCL-R total score alone, they may have missed important nuances that may have made for a better informed and arguably more ethical recommendation.

Potential Clinical Utility of Subtyping Study Results

Research has consistently shown that offenders falling within different psychopathy measure subtype groups differ on clinically meaningful outcome variables, such as treatment progress, violence risk, and types of psychopathology. Indeed, many of these variables have important value in the context of high-stakes criminal justice considerations. For example, the finding that individuals in some subtypes are considered to show better treatment progress than others (Klein Haneveld et al., 2018) could potentially affect sentencing for some offenders.

For LPA subtype correlate findings to be useful for clinical practice, however, clinicians scoring the PCL-R would need to be able to reliably assign offenders to their correct subgroups so that they could come to appropriate conclusions about elements such as future violence. Currently, none of the many existing subtype studies have examined whether anyone (e.g., researchers, clinicians, research volunteers) can identify the correct latent class for a sample of PCL-R profiles (e.g., Hare et al., 2018; Klein

Haneveld et al., 2018; Krstic et al., 2018; Lehmann et al., 2019; McCallum et al., 2020; Neumann et al., 2016). If practitioners cannot make these assignments with a reasonable degree of accuracy, the LPA results and their suggested correlates from published studies cannot be directly useful for real-world, high-stakes decision-making.

Because the PCL-R facets are not all based on the same number of items, researchers usually present psychopathy LPA results in the metric of mean item score for each facet (see Figure 1). These mean item scores can range from 0.00 to 2.00 for each facet. Although this is helpful when comparing subgroups to one another and differences between facets within these subgroups, it does present one potential challenge for clinical application. Standard PCL-R scoring procedures lead clinicians to sum item scores on each facet, providing summed total scores for each facet. Because each facet does not have the same number of items (Facets 1 and 2 have four items each, whereas Facets 3 and 4 have five items each), the range of possible summed scores differs among the four facets (i.e., 0.00 to 8.00 for Facets 1 and 2; 0.00 to 10.00 for Facets 3 and 4). Therefore, one challenge in translating LPA findings to practice is whether evaluators can look at a PCL-R profile's pattern of *summed facet scores* in the field, compare those scores to LPA research findings that are presented using *mean item scores*, and correctly classify that PCL-R profile into the correct subgroup.

The Current Study

The current study used PCL-R scores and LPA-based subtype classifications from an already completed study (N = 615; McCallum et al., 2020) to examine the ability of graduate students and forensic psychologists to accurately classify offenders into their LPA-identified subtype groups. The LPA results identified four subtype groups:

prototypic, callous-conning, sociopathic, and non-psychopathic/general offenders (McCallum et al., 2020). I selected 30 offenders from each subtype group and used PCL-R scores from this sample of 120 offenders in this two-study project. Participants in both studies were asked to classify each profile into one of four subtype groups (prototypical, callous-conning, sociopathic, general offender) and rate their level of confidence in each classification.

There were two parts to this project: a pilot study of trained clinical psychology graduate students (Study 1) and a subsequent study with forensic clinicians (Study 2). In Study 1, the students completed the classification task twice (order counterbalanced), meaning they were asked to see and classify each offender's profile twice. Throughout both classification conditions, participants were provided four graphs depicting the PCL-R subtypes from the McCallum et al. (2020) study. However, in one condition, they were provided with the offender's mean item scores and PCL-R total score and in the other condition they were provided with the offender's summed facet scores and PCL-R total score. The condition asking for classifications based on mean item scores best represents how clinicians would have to make classifications based on the current published research literature, whereas the condition asking for classification based on summed facet scores best represents the information clinicians commonly use to describe and interpret PCL-R results in the field.

I hypothesized that it would be more difficult to make classifications in the summed facet score condition (i.e., comparing summed scores for offenders to mean item scores on the graphs from the McCallum study) and that there would be a low level of accuracy in this condition. Using the two study conditions allowed for the examination of whether the manner in which LPA findings are currently presented in the research literature (i.e., mean item scores) were easily translated to clinical practice or whether it would be appropriate for researchers to begin to present their findings in a way that better matches how the PCL-R is used in the field (summed facet scores). I also used Study 1 findings to inform the methodology for Study 2 with clinicians.

In Study 2, psychologists with experience using the PCL-R were also asked to complete a profile classification task. They were presented with the four graphs from Study 1 depicting the four subtype groups found by McCallum and colleagues (2020) but were asked to make classifications after receiving a completed PCL-R score sheet (as opposed to only facet and total scores). Specifically, participants in Study 2 were presented with item, facet, and total scores for the offenders, similar to what they would use to assign and interpret in the field. Using this type of presentation method increased the ecological validity of the classification task (i.e., comparing a complete set of PCL-R scores to research generated figure).

For both studies, I examined overall level of accuracy (0%-100%) for each student/clinician and the average level of accuracy across students/clinicians. I also examined whether accuracy was similar for each subtype and attempted to identify the characteristics of profiles that were most- and least-likely to be correctly classified. Finally, I investigated whether participants' confidence in the accuracy of their classifications was useful for providing information about when a profile classification was likely to be inaccurate. I hypothesized that, if these confidence ratings are useful, I would find that participants reported significantly higher levels of confidence for correct classifications than incorrect classifications.

I expected that students and clinicians would be best at correctly classifying offenders falling into the general offender subtype group because it should be straightforward to identify offenders with low scores across all facets. I expected that the classification task would be more difficult for the other subtype groups because they are based on multiple facet elevations. Although it may be straightforward to classify some prototypic offenders with high scores across all facets, it may also be difficult to determine if those with less uniform elevations belong in the prototypic subtype or one of the subtypes with varied elevations (i.e., callous-conning, sociopathic).

Overall, because some, if not many, profiles may be difficult to classify, I expected only a modest level of classification accuracy (≈ 60%). I expected students and clinicians to recognize when profiles were especially difficult to classify (i.e., do not obviously fit into a subtype group), which should have led them to have lower confidence ratings for more difficult profiles, as well as lower confidence ratings for incorrectly classified profiles. Therefore, because I expected to find higher classification accuracy for the general offender subtype group, I also expected to find participants to provide higher confidence ratings when classifying these profiles.

CHAPTER II

Study 1

Method

Participants

Participants were 13 clinical psychology doctoral students at Sam Houston State University (SHSU) who had been trained in forensic assessment and had at least one year of clinical forensic training experience. Participants had an average of 3.46 years in the doctoral program (*SD* = 1.56). Most of the participants were female (*n* = 11; 84.62%) and White/European American (*n* = 12; 92.31%). All participants were recruited via email and received a $50 Amazon gift card for their participation.

Although 13 individuals participated in the study, one was eliminated due to concerns that they did not understand the instructions based on especially low accuracy. Therefore, the final sample of participants consisted of 12 clinical psychology doctoral students. Five participants (41.67%) had experience using the PCL-R in supervised clinical practice and seven participants (58.33%) had previously taken a psychopathy course. Most participants (71.4%) who endorsed previously taking a psychopathy course took a course by the same instructor at SHSU. Three of the five students who had experience using the PCL-R in clinical practice reported having completed a psychopathy course.

Materials

Training Materials. I created a PowerPoint presentation consisting of information regarding psychopathy (six slides), the PCL-R (two slides), research relevant to psychopathy subtyping (11 slides), and the participants' role in the current study (three slides). More specifically, I presented on a brief history of the psychopathy construct, empirical models, and measurement. Further, I also presented on information regarding the development, underlying structure, and the clinical utility of the PCL-R. Additionally, I provided subtyping information from relevant academic and clinical texts (i.e., Neumann et al., 2016 and Hicks & Drislane, 2018), as well as empirical research. Special focus was on McCallum and colleagues' (2020) LPA findings.

Two training sessions were held to accommodate scheduling conflicts and availability of the participants. Each training session lasted approximately three hours and was provided via Zoom. The training included a presentation of the aforementioned information, a question-and-answer period about the information presented, and a description of the classification task. Participants were then instructed to complete a practice classification task comprised of 12 profiles. Once everyone in the training competed the practice, I reviewed the responses to see if there were any major issues to address. The purpose of having participants complete these practice classifications in the training was to ensure familiarity with the classification task, as opposed to collecting data for a rater agreement study. Just as evaluators in the field would not be required to be trained to a certain level of reliability before performing this type of task, I did not require the Study 1 participants to obtain any specified level of accuracy during the training. Nevertheless, participants performed very well on the practice cases, with an overall accuracy rate of 94.9% when data were presented as mean item scores and 85.9% when data were presented as summed facet scores.

Neither training presented any issues with the practice classification task. All participants completed the task without any technical or conceptual problems.

Furthermore, after I reviewed practice responses and everyone returned to the Zoom room for a subsequent question-and-answer period, no participants expressed confusion regarding the task and all clarifying questions suggested understanding of the task. Once all questions were answered, the training ended, and the participants were emailed a link to the Study 1 classification task.

Demographics. Participants completed a brief questionnaire that included questions regarding demographics, prior experience using the PCL-R and prior psychopathy courses taken.

PCL-R Profiles. McCallum and colleagues (2020) performed an LPA of PCL-R scores collected from 615 sexually violent predator (SVP) evaluations (see Boccaccini et al., 2008; Boccaccini et al., 2014). Based on this analysis, each of the 615 evaluated individuals were placed into one of four subtypes: prototypic, callous-conning, sociopathic, and non-psychopathic/general offenders. The prototypic subgroup consisted of 38.9% of the sample (*n* = 239), exhibited an average PCL-R total score of 27.22 (*SD* = 3.59), and displayed the highest elevations across all facets (see Figure 1). Although the callous-conning group, consisting of 25% of the sample (*n* = 154), and the sociopathic group, consisting of 15.6% of the sample (*n* = 96), displayed similar mean PCL-R total scores (*M* = 18.12 and 18.22, respectively), facet level scores differentiated the two subgroups. Specifically, the callous-conning group had higher scores on Facets 1 and 2, and lower scores on Facets 3 and 4 compared to the sociopathic group (see Figure 1). Lastly, the subgroup of non-psychopathic, or general offenders, consisted of 20.5% of the sample (*n* = 126), displayed low elevations across all facets, and exhibited the lowest PCL-R total score of 9.93 (*SD* = 3.25; see Figure 1). Although there were 615 PCL-R profiles utilized for the LPA, the current study only used subtype information for 120 of these evaluated individuals.

I used posterior class assignment probabilities from the McCallum et al. (2020) LPA analyses to choose 30 offenders from each subtype group. For each subgroup, the probability values for each class were sorted from highest to lowest. Then, 10 individuals who had a relatively high probability value for the class (i.e., from the top third), 10 individuals who had a relatively low probability value for the class (i.e., from the bottom third), and 10 individuals with probability values from the middle-third were randomly selected. This process was repeated for each subgroup, yielding 30 offenders from each subgroup. Therefore, among the 120 evaluated individuals chosen for this study, 30 were from the prototypic subgroup, 30 from the callous-conning subgroup, 30 from the sociopathic subgroup, and 30 from the general offender subgroup. Within each subgroup, there were 10 high, 10 middle, and 10 low probability profiles.

Each of the 120 PCL-R profiles were presented in two ways. For one set of PCL-R profiles, participants were provided graphical representations of the prototypical facet scores of each of the four subgroups found by McCallum and colleagues (2020). Below these graphs, the participants were provided a table containing the total score and mean item scores for each of the four PCL-R facets for one individual (Appendix A). The other set of PCL-R profiles were also presented with the same graphical representations of the prototypical facet scores of each of the four subgroups found by McCallum and colleagues (2020); however, for these profiles, below these graphs, the participants were provided a table containing the total score and summed facet scores for each of the four PCL-R facets for one individual (Appendix B).

For both conditions, participants were asked to classify each profile into one of the four subtype groups. Specifically, they were asked to check the box of the subtype they believe best represented that particular profile. Participants were randomly assigned to complete either the mean item or summed facet profiles first. They completed all profile classifications (presented in a randomly assigned order) in this condition before starting the next condition (where the profiles in that condition were then presented in a randomly assigned order).

Confidence Rating. After each classification, participants were asked to rate how confident they were in their classification on a scale from 1 to 10 (1 = not confident at all, 10 = extremely confident).

Open-Ended Questions. Participants were asked questions to better assess the manner with which they completed the classification task. More specifically, after participants completed all the classifications in the first condition, they were asked to describe the methods they used to assign each profile to a subtype. After participants completed all the classifications in the second condition, they were once again asked to report how they assigned each profile to a subtype in as much detail as possible. Participants were then also asked which presentation method was most helpful in classifying the scores they were provided (mean item score, summed facet score), what other information could have been provided to make the task easier, if the task would have been easier if they had been provided a graphical representation of each individual's score in addition to a table of scores, and any other information they would like to share about the task.

Results

Overall Classification Accuracy

I used descriptive statistics (i.e., percentage values, means) to examine classification accuracy for each participant across presentation methods and subtypes. Table 1 provides classification accuracy information separately for each of the 12 participants, for each method (i.e., mean item, summed facet), and for each subtype.

For overall accuracy, which I defined as the correct classification rate across all 120 profiles, participants had similar levels of classification accuracy when provided mean item scores (M = 71.7%) and summed facet scores (M = 69.1%; see Table 1). I used a dependent samples t-test to examine if there was a statistically significant difference in classification accuracy between the presentation types. This difference was not large enough to reach statistical significance in this small sample [$t(11) = 1.91$, p = .083], although it was moderate in size (Cohen's d =.55).

Overall accuracy was 67.83% for mean item score profiles and 66.83% for summed facet score profiles for the five participants who had experience using the PCL-R. Overall accuracy was 71.43% for mean item score profiles and 68.21% for summed facet score profiles for the seven participants who endorsed being trained to use the PCL-R. I used an independent samples t-tests to see if these two groups differed on their classification accuracy for each presentation method. There was not a statistically significant difference in classification accuracy between those who had experience using the PCL-R and those who did not when presented with mean item profiles [$t(10) = -1.738$, p = .113; Cohen's d = .97] and when presented with summed facet profiles [$t(10) = -.727$, p = .484; Cohen's d = .42]. There was also not a statistically significant difference in classification accuracy between those who had prior PCL-R training and those who did not when presented with mean item scores [$t(10) = -.145$, $p = .891$] and when presented with summed facet scores [$t(10) = -.341$, $p = .705$].

Accuracy for individual evaluators ranged from 55.0% to 81.7% for mean item scores, and 51.7% to 85.0% for summed facet scores (see Table 1). To perform better than chance (i.e., 25.0%), participants had to correctly classify at least 39 of the 120 profiles (i.e., $\geq$ 32.5%). Binomial tests revealed each participant performed better than chance on each presentation method at $p < .001$.

Classification Accuracy for Subtypes

Table 1 also provides classification accuracy findings for each psychopathy subtype. In terms of absolute value, classification accuracy was highest for general offender profiles. Specifically, the average level of accuracy across evaluators for general offenders was 81.11% ($SD = 16.66$) for mean item scores and 76.39% ($SD = 18.17$) for summed facet scores. Classification accuracy was also relatively strong for sociopathic offender profiles. Specifically, the average level of accuracy across evaluators for sociopathic profiles was 70.83% ($SD = 16.76$) for mean item scores and 80.00% ($SD = 12.47$) for summed facet scores.

Classification accuracy was, however, somewhat lower for prototypic and callous-conning profiles. For prototypic offenders, the average level of accuracy across evaluators was 54.72% ($SD = 18.34$) for mean item scores and 52.22% ($SD = 24.63$) for summed facet scores. For callous-conning offenders, the average level of accuracy across evaluators was 80.28% ($SD = 9.48$) for mean item scores and 67.78% ($SD = 24.54$) for summed facet scores.

Table 1

Doctoral Student Participants' Classification Accuracy for each Subtype and each Presentation Method

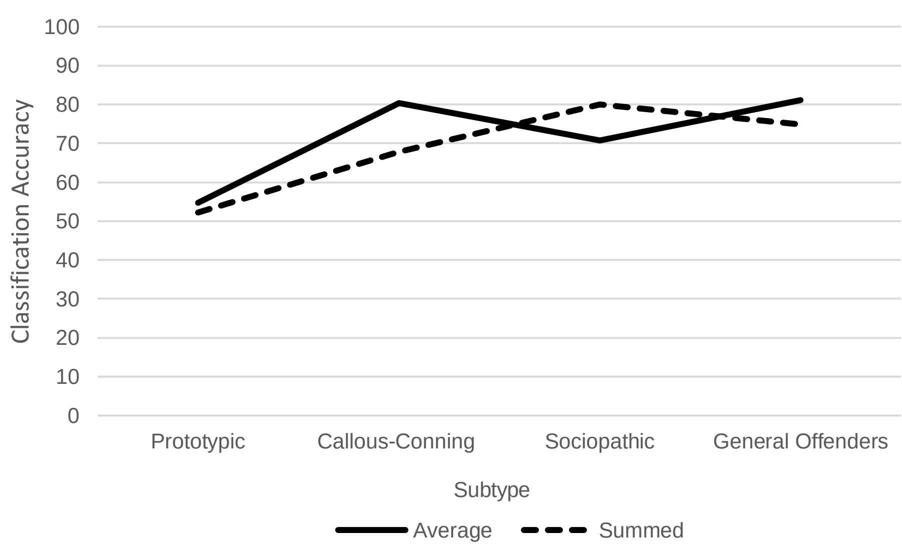
Participant	Presentation Type	General Offender	Sociopathic	Prototypic	Callous-conning	Overall Accuracy
1	Mean	93.33	83.33	63.33	86.67	81.67
	Sum	76.67	100.00	70.00	93.33	85.00
2	Mean	76.67	83.33	73.33	70.00	75.83
	Sum	76.67	86.67	56.67	63.33	70.83
3	Mean	96.67	66.67	43.33	73.33	70.00
	Sum	100.00	76.67	53.33	50.00	70.00
4	Mean	73.33	83.33	53.33	86.67	74.17
	Sum	40.00	86.67	30.00	93.33	62.50
5	Mean	50.00	60.00	33.33	76.67	55.00
	Sum	46.67	70.00	13.33	76.67	51.67
6	Mean	73.33	80.00	33.33	86.67	68.33
	Sum	70.00	73.33	20.00	93.33	64.17
7	Mean	90.00	80.00	36.67	100.00	76.67
	Sum	83.33	63.33	33.33	93.33	68.33
8	Mean	63.33	36.67	83.33	76.67	65.00
	Sum	86.67	63.33	83.33	30.00	65.83
9	Mean	63.33	90.00	76.67	90.00	80.00
	Sum	73.33	96.67	80.00	70.00	80.00
10	Mean	100.00	80.00	33.33	73.33	71.67
	Sum	100.00	73.33	50.00	70.00	73.33
12	Mean	96.67	60.00	63.33	73.33	73.33
	Sum	76.67	93.33	86.67	46.67	75.83
13	Mean	96.67	46.67	63.33	70.00	69.17
	Sum	86.67	76.67	50.00	33.33	61.67
Mean (SD)	Mean	81.11 (16.66)	70.83 (16.76)	54.72 (18.34)	80.28 (9.48)	71.74 (7.16)
	Sum	76.39 (18.17)	80.00 (12.47)	52.22 (24.63)	67.78 (23.54)	69.10 (8.93)

Note. Mean = mean item score profiles; Sum = summed facet score profiles

I used a 2 (presentation method) x 4 (subtype) repeated MANOVA to investigate if there were statistically significant differences in accuracy rates among subtypes and presentation methods. Figure 2 provides a graphical depiction of the results of that analysis. There was a main effect of subtype on accuracy $F(3, 9) = 5.91$, $p = .02$, $\eta p^2 = .66$. This effect indicated that classification accuracy was lower for prototypic offenders ($M = 53.47$, $SE = 5.88$) than general ($M = 78.75$, $SE = 4.57$), sociopathic ($M = 75.42$, $SE = 3.64$), and callous conning ($M = 74.03$, $SE = 4.49$) offenders.

Figure 2

Doctoral Student Participants' Classification Accuracy for each Subtype and each Presentation Method



The main effect of subtype on accuracy between the two presentation methods was not statistically significant $F(1, 11) = 3.64$, $p = .08$, $\eta p^2 = .25$. In terms of absolute value, accuracy was higher for mean item scores ($M = 71.74$, $SE = 2.07$) than summed facet scores (*M* = 69.10, *SE* = 2.58). The two way interaction was not statistically significant $F(3, 9) = 2.20$, $p = .16$, although there was a moderate effect size of $\eta p^2 = .42$.

The overall pattern was that participants had a higher classification accuracy for the callous-conning, general offender, and prototypic subtypes when presented as mean item scores and a higher classification accuracy for the sociopathic subtype when presented as summed facet scores. The largest differences were as follows: between the general offender subtype (*M* = 81.11, *SD* = 16.66) and prototypic subtype (*M* = 54.72, *SD* = 18.34) when presented as mean item scores (Cohen's *d* = 1.51), between the prototypic subtype and callous-conning subtype (*M* = 80.28, *SD* = 9.48) when presented as mean item scores (Cohen's *d* = 1.75), and between the sociopathic subtype (*M* = 80, *SD* = 12.47) and prototypic subtype (*M* = 52.22, *SD* = 24.63) when presented as summed facet scores (Cohen's *d* = 1.42).

Was Classification for Subtypes Better than Chance?

I used binomial tests to investigate whether participants performed better than chance on each subtype and across each presentation method. Table 2 includes the *p*-values for these tests, which I calculated separately for each participant, for each subtype. A statistically significant *p*-value indicated that classification was significantly better than chance.

Although many participants performed better than chance on each subtype and across each presentation at $p < .001$, some participants performed no better than chance for some subtypes and presentations ($p = .05$ to .96; Table 2). Table 2 also shows some notable differences. For example, more participants performed no better than chance for the prototypic subtype both presented as mean item scores (n = 4) and summed facet scores (n = 4), than any other subtype.

Table 2

Binomial Test p-values for Doctoral Student Participants across Subtype and Presentation Type

Part	Presentation Type	General Offender	Sociopathic	Prototypic	Callous-conning
1	Mean	<.001	<.001	<.001	<.001
	Sum	<.001	<.001	<.001	<.001
2	Mean	<.001	<.001	<.001	<.001
	Sum	<.001	<.001	<.001	<.001
3	Mean	<.001	<.001	.022	<.001
	Sum	<.001	<.001	<.001	.003
4	Mean	<.001	<.001	<.001	<.001
	Sum	.051	<.001	.326	<.001
5	Mean	.003	<.001	.197	<.001
	Sum	.008	<.001	.963	<.001
6	Mean	<.001	<.001	.197	<.001
	Sum	<.001	<.001	.798	<.001
7	Mean	<.001	<.001	.106	<.001
	Sum	<.001	<.001	.197	<.001
8	Mean	<.001	.106	<.001	<.001
	Sum	<.001	<.001	<.001	.326
9	Mean	<.001	<.001	<.001	<.001
	Sum	<.001	<.001	<.001	<.001
10	Mean	<.001	<.001	.197	<.001
	Sum	<.001	<.001	.003	<.001
12	Mean	<.001	<.001	<.001	<.001
	Sum	<.001	<.001	<.001	.008
13	Mean	<.001	.008	<.001	<.001
	Sum	<.001	<.001	.003	.197

Note. Part = Participant; Mean = mean item score profiles; Sum = summed facet score profiles

Differences in Subtype Accuracy for Individual Participants

I examined differences in accuracy between presentation methods for individual participants using McNemar's test (Table 3). Descriptive statistics and results from these analyses are provided in Table 3. Out of the 12 participants, only one had a significant difference between presentation methods. Specifically, Participant 4 accurately classified 74.17% of profiles presented as mean item scores and 62.50% of profiles presented as summed facet scores, which was statistically significant, $p = .01$.

Table 3

McNemar's Test and Odds Ratio for Doctoral Student Participants' Accuracy

Participant	% correct for mean item	% correct for summed facet	McNemar Test Statistic	*p*-value	Odds Ratio
1	81.67	85.00	0.56	0.45	0.60
2	75.83	70.83	0.83	0.36	1.50
3	70.00	70.00	0.05	0.82	1.00
4	74.17	62.50	7.04	0.01	3.80
5	55.00	51.67	0.45	0.50	1.50
6	68.33	64.17	0.70	0.40	1.56
7	76.67	68.33	3.68	0.06	2.67
8	65.00	65.83	0.00	1.00	0.95
9	80.00	80.00	0.05	0.83	1.00
10	71.67	73.33	0.13	0.72	0.60
12	73.33	75.83	0.11	0.74	0.84
13	69.17	61.67	1.73	0.19	1.64
Mean					1.47

Classification Accuracy and Posterior Class Assignment Probability

There were 30 profiles for each subtype. These 30 profiles varied across posterior class assignment probabilities, with 10 high probability profiles, 10 low probability profiles, and 10 profiles with mid-range class probabilities. Tables 4, 5, 6, and 7 includes the descriptive values for participants' classification accuracy for each subtype and each presentation method by posterior class assignment probabilities.

I used bivariate correlations to examine the association between classification accuracy and posterior class probability across participants. Specifically, I correlated the overall classification accuracy for each subtype with the posterior class probability for each subtype for both presentation methods.

Table 4

Doctoral Student Participants' Classification Accuracy Overall and for the General Offender Subtype across Presentation Method, by Posterior Class Assignment Probabilities

Part.	Pres Type	Overall	Gen High	Gen Med	Gen Low
1	M	81.67	100	100	80
	Sum	85.00	100	90	40
2	M	75.83	100	80	50
	Sum	70.83	100	80	50
3	M	70.00	100	100	90
	Sum	70.00	100	100	100
4	M	74.17	100	70	50
	Sum	62.50	70	30	20
5	M	55.00	60	70	20
	Sum	51.67	70	40	30
6	M	68.33	90	80	50
	Sum	64.17	100	70	40
7	M	76.67	100	100	70
	Sum	68.33	100	90	60
8	M	65.00	100	60	30
	Sum	65.83	100	100	60
9	M	80.00	90	60	40
	Sum	80.00	90	80	50
10	M	71.67	100	100	100
	Sum	73.33	100	100	100
12	M	73.33	100	90	100
	Sum	75.83	100	80	50
13	M	69.17	100	100	90
	Sum	61.67	100	90	70
M (SD)	M	71.74 (7.16)	95 (11.68)	84.17 (16.21)	64.17 (27.78)
	Sum	69.10 (8.93)	94.17 (11.65)	79.17 (22.75)	55.83 (24.66)

Note. M = mean item score profiles; Sum = summed facet score profiles; Gen High = General offender subtype with high posterior class assignment probability; Gen Med = General offender subtype with medium posterior class assignment probability; Gen Low = General offender subtype with low posterior class assignment probability

Table 5

Doctoral Student Participants' Classification Accuracy for the Sociopathic Subtype across Presentation Method, by Posterior Class Assignment Probabilities

Part.	Pres Type	Soc High	Soc Med	Soc Low
1	M	100	90	60
	Sum	100	100	100
2	M	90	90	70
	Sum	90	70	100
3	M	90	80	30
	Sum	100	80	50
4	M	90	100	60
	Sum	100	100	60
5	M	80	80	20
	Sum	90	80	40
6	M	100	80	60
	Sum	90	90	40
7	M	100	90	50
	Sum	100	70	20
8	M	60	30	20
	Sum	90	80	20
9	M	100	90	80
	Sum	100	100	90
10	M	100	90	50
	Sum	100	80	40
12	M	90	70	20
	Sum	100	100	80
13	M	90	50	0
	Sum	100	90	40
M (SD)	M	90.83 (11.65)	78.33 (19.92)	43.33 (24.62)
	Sum	96.67 (4.92)	86.67 (11.55)	56.67 (29.02)

Note. M = mean item score profiles; Sum = summed facet score profiles; Soc High = Sociopathic offender subtype with high posterior class assignment probability; Soc Med = Sociopathic offender subtype with medium posterior class assignment probability; Soc Low = Sociopathic offender subtype with low posterior class assignment probability

Table 6

Doctoral Student Participants' Classification Accuracy for the Prototypic Subtype across Presentation Method, by Posterior Class Assignment Probabilities

Part.	Pres Type	Pro High	Pro Med	Pro Low
1	M	100	70	20
	Sum	100	80	30
2	M	90	70	60
	Sum	80	50	40
3	M	100	30	0
	Sum	100	60	0
4	M	80	50	30
	Sum	40	30	20
5	M	60	20	20
	Sum	40	0	0
6	M	40	30	30
	Sum	40	10	10
7	M	80	30	0
	Sum	80	20	0
8	M	90	70	90
	Sum	90	90	70
9	M	100	90	40
	Sum	100	80	60
10	M	90	10	0
	Sum	100	40	10
12	M	100	70	20
	Sum	100	90	70
13	M	90	50	50
	Sum	70	50	30
M (SD)	M	85 (18.34)	49.17 (25.03)	30 (26.97)
	Sum	78.33 (25.17)	50 (31.04)	28.33 (26.57)

Note. M = mean item score profiles; Sum = summed facet score profiles; Pro High = Prototypic offender subtype with high posterior class assignment probability; Pro Med = Prototypic offender subtype with medium posterior class assignment probability; Pro Low = Prototypic offender subtype with low posterior class assignment probability

Table 7

Doctoral Student Participants' Classification Accuracy for the Callous-conning Subtype across Presentation Method, by Posterior Class Assignment Probabilities

Part.	Pres Type	Cal High	Cal Med	Cal Low
1	M	100	90	70
	Sum	100	80	100
2	M	100	70	40
	Sum	100	60	30
3	M	80	80	60
	Sum	60	50	40
4	M	100	90	70
	Sum	100	100	80
5	M	100	80	50
	Sum	100	90	40
6	M	90	100	70
	Sum	100	100	80
7	M	100	100	100
	Sum	100	100	80
8	M	100	80	50
	Sum	70	10	10
9	M	100	80	90
	Sum	90	70	50
10	M	80	80	60
	Sum	80	80	50
12	M	70	80	70
	Sum	70	50	20
13	M	90	70	50
	Sum	40	40	20
M (SD)	M	92.50 (10.55)	83.33 (9.85)	65 (17.32)
	Sum	84.17 (20.21)	69.17 (28.11)	50 (28.92)

Note. M = mean item score profiles; Sum = summed facet score profiles; Cal High = Callous-conning offender subtype with high posterior class assignment probability; Cal Med = Callous-conning offender subtype with medium posterior class assignment probability; Cal Low = Callous-conning offender subtype with low posterior class assignment probability

Table 8 provides the correlation and *p*-values for each participant across each subtype broken down by class probability (i.e., high, medium, low). Within each subtype and presentation method, classification accuracy was significantly correlated (p =.00) with class probability, and these correlations were moderate to strong in size (r = .65 to

.82). In other words, the higher the posterior class assignment probability of the profile, the more accurately it was placed into the correct subtype by participants.

Table 8

Correlations between Accuracy and Class Probability across Doctoral Student Participants

Presentation Method	Correlation	*p*-value
Mean Item		
General Offender	0.80	0.00
Sociopathic	0.74	0.00
Prototypic	0.65	0.00
Callous-conning	0.71	0.00
Summed Facet		
General Offender	0.78	0.00
Sociopathic	0.82	0.00
Prototypic	0.71	0.00
Callous-conning	0.69	0.00

Confidence and Accuracy

I used point-biserial correlations to examine whether there was a relation between accuracy and confidence. I calculated these correlations separately for each participant. Table 9 provides the correlation and *p*-values for these analyses. Confidence and classification accuracy were significantly correlated in that participants had more confidence when their classification was accurate than when it was inaccurate (see Table 9). Specifically, for profiles presented as mean item scores, correlations between confidence and accuracy ranged from 0.19 to 0.36 across participants (M = 0.28) and for profiles presented as summed facet scores, correlations between confidence and accuracy ranged from 0.21 to 0.34 across participants (M = 0.25).

Table 9

Point-biserial Correlations between Accuracy and Confidence Rating for each Doctoral Student Participant

Participant	Mean Item		Summed Facet	
	r	*p*	r	*p*
1	0.34	0.00	0.25	0.01
2	0.19	0.04	0.22	0.02
3	0.33	0.00	0.27	0.00
4	0.26	0.01	0.24	0.01
5	0.29	0.00	0.25	0.01
6	0.27	0.00	0.24	0.01
7	0.20	0.03	0.25	0.01
8	0.22	0.02	0.28	0.00
9	0.32	0.00	0.34	0.00
10	0.25	0.01	0.32	0.00
12	0.36	0.00	0.21	0.02
13	0.31	0.00	0.21	0.02
Mean (SD)	0.28	--	0.25	--

I also used t-tests to compare confidence ratings between each participant's accurate and inaccurate classifications. Table 10 provides the descriptive statistics and results for each participant for the mean item presentation method and Table 11 provides the descriptive statistics and results for each participant for the summed facet presentation method. There was a significant difference ($p < .05$) for each participant. In other words, each participant's mean confidence rating was higher for accurate classifications than inaccurate classifications (Cohen's d = .44 to .94, M = 0.65 and Cohen's d = .45 to .91, M = 0.62 for mean item scores and summed facet scores respectively).

Table 10

T-tests between Accuracy and Confidence Rating for Mean Item Profiles for Doctoral Student Participants

Mean Item						
Participant	Mean (SD) confidence for accurate classifications	N	Mean (SD) confidence for inaccurate classifications	N	*p*-value	Cohen's *d*
1	7.54 (2.30)	97	5.38 (2.22)	21	0.00	0.94
2	8.11 (1.07)	91	7.62 (1.21)	29	0.04	0.44
3	4.77 (2.37)	83	3.19 (1.33)	36	0.00	0.75
4	7.69 (1.59)	89	6.74 (1.50)	31	0.01	0.61
5	5.44 (1.87)	66	4.44 (1.41)	54	0.00	0.60
6	7.16 (1.23)	82	6.34 (1.42)	38	0.00	0.63
7	5.65 (1.92)	92	4.79 (1.29)	28	0.01	0.48
8	5.04 (1.96)	78	4.19 (1.55)	42	0.01	0.47
9	7.01 (1.76)	96	5.58 (1.44)	24	0.00	0.84
10	7.41 (2.60)	86	5.79 (3.38)	34	0.02	0.57
12	7.07 (1.61)	88	5.66 (1.77)	32	0.00	0.85
13	7.17 (2.06)	82	5.72 (2.05)	36	0.00	0.70
Mean	6.67		5.45			0.65

Table 11

T-tests between Accuracy and Confidence Rating for Summed Facet Profiles for Doctoral Student Participants

Summed Facet						
Participant	Mean (SD) confidence for accurate classifications	N	Mean (SD) confidence for inaccurate classifications	N	*p*-value	Cohen's *d*
1	7.30 (0.66)	96	6.00 (1.29)	13	0.01	0.80
2	8.12 (0.83)	85	7.71 (0.67)	35	0.02	0.52
3	4.82 (2.66)	84	3.41 (1.18)	34	0.00	0.60
4	7.57 (1.47)	75	6.78 (1.78)	45	0.01	0.50
5	4.87 (1.42)	62	4.1 (1.53)	58	0.01	0.52
6	6.70 (1.10)	77	6.14 (1.10)	43	0.01	0.51
7	5.23 (2.22)	82	4.13 (1.46)	38	0.00	0.55
8	5.44 (1.91)	79	4.37 (1.39)	41	0.00	0.61
9	7.07 (1.93)	96	5.33 (1.88)	24	0.00	0.91
10	6.41 (2.55)	88	4.44 (2.61)	32	0.00	0.77
12	6.90 (1.54)	91	6.1 (1.74)	29	0.02	0.50
13	7.49 (1.94)	74	6.65 (1.74)	46	0.02	0.45
Mean	6.49		5.43			0.62

Summary of Open-Ended Responses

Responses to the qualitative questions indicated that some participants (n = 4) believed that the mean item score profiles were easier to classify. Participants described a diverse range of methods they used to classify the mean item profiles, sometimes using multiple methods in succession. For profiles presented as mean item scores, five participants (41.67%) considered the total score of the profile and then considered the pattern of the facet scores, whereas three participants (25%) considered the patterns of the facets and then considered the total score of the profile. One participant described using both the total score and the pattern of facet scores in tandem, one participant (8.33%) considered the pattern of the facet scores and only sometimes considered total score, and two participants (16.67%) only reported considering the pattern of the facet scores with no mention of considering total scores.

The participants' classification methods for the profiles presented as summed facet scores were also mixed. Three participants (25%) considered total score and then the pattern of the facets, whereas three participants (25%) considered the pattern of the facet scores and then considered total score. Four participants (33.33%) converted the summed facet scores into mean item scores, with two of those participants then using the patterns of those new values to help them classify the profile (one of whom also drew their own graphs if necessary), one participant considering the patterns of the new values and then total score, and one participant considering the total score and then the pattern of the facet scores they calculated. One participant (8.33%) only reported using total scores and one participant (8.33%) reported using both the pattern of the facet scores provided and the total scores in tandem.

I investigated the classification methods for the two participants (participants 1 and 9) who more accurately classified profiles compared to the other participants. Although these participants used different strategies for the profiles presented as mean item scores, they used similar strategies for profiles presented as summed facet scores. Specifically, both participants converted the summed facet score values into mean item scores before proceeding. Interestingly, however, one participant (participant 6) who did not perform as well also used mental division to calculate the mean item score rather than using the summed facet scores provided. Furthermore, when comparing participants who less accurately classified profiles compared to the other participants, there was no consistent strategy utilized.

Discussion

This study was a first step in investigating the clinical utility of psychopathy subtypes using the PCL-R. It provides encouraging preliminary evidence that it may be plausible to translate findings from sophisticated psychopathy subtyping research to applied clinical practice, as participants were able to classify individual profiles into their correct empirically derived subtypes with a high degree of accuracy. However, it also revealed that some characteristics of PCL-R profiles may engender more utility than others based on the ease at which an individual is able to accurately identify the profile's empirical subtype.

Overall, the doctoral student participants in this study performed better than hypothesized with classifying PCL-R profiles into their correct subtype. The difference in classification accuracy between the two presentation methods was not statistically significant, although it was moderate in effect size with participants performing better when provided mean item scores compared to summed facet scores. This indicates the way subtyping results are presented in the literature (i.e., as mean item scores) may be a hindrance when seeking to apply this approach in real-world settings, as forensic evaluators are typically presented with summed facet scores instead. As hypothesized, classification accuracy was highest for the general offender subtype; however, only when presented as mean item scores. Interestingly, classification accuracy was highest for the sociopathic subtype when the profiles were presented as summed facet scores and the sociopathic subtype was the only subtype for which participants performed better on the summed facet presentation method compared to the mean item presentation method.

As hypothesized, participants performed worse when classifying the prototypic subtype, with 33.33% of participants performing no better than chance for this subtype. This finding is concerning, as prototypic offenders score higher on violence risk measures (Lehmann et al., 2019; McCallum et al., 2020; Neumann et al., 2016), display more features of personality disorders (Klein Haneveld et al., 2018; McCallum et al., 2020), have committed more violent offenses (Kristic et al., 2018; Lehmann et al., 2018), and drop out of treatment at a significantly higher rate (Klein Haneveld et al., 2018). It could be that prototypic profiles are more difficult to classify because the subtype subsumes primary and secondary variants of psychopathy and is not homogeneous. Further research is needed to better assess those difficulties.

Classification accuracy was significantly correlated with posterior class probability within each subtype and across presentation methods. Although this correlation supports the usefulness of these probabilities, it draws concern to those profiles which do not represent a “typical” profile for that subtype. Additionally, participants accurately appraised their own ability to perform this task. Doctoral student participants expressed more confidence with profiles they accurately classified compared to profiles they inaccurately classified. Participants reported various methods of classification and there was no particular method that appeared to be more or less helpful as evaluated by classification accuracy. Of note, however, some participants reported that the profiles presented as mean item scores were easier to classify than profiles presented as summed facet scores. Although there was not a statistically significant difference in classification accuracy between presentation type, the moderate effect size coupled with some participants reporting that the mean item profiles were easier to classify suggests that it may be more useful for researchers to present subtyping findings in a manner that is consistent with clinical practice.

The biggest limitation of this study was that the participants consisted of students that, although forensically trained, had limited to no experience using the PCL-R in field settings. I attempted to mitigate this by having the participants complete a three-hour training and conduct practice scoring prior to completing the classification task. Further, the sample size of this study was small, which can affect statistical findings and limit the generalizability of the findings. Lastly, participants were only provided a table of facet scores and a total score for each profile presented; the item level data was not provided. This limited the information participants could potentially use to aid in classification.

Study 2 was aimed to address some of these limitations. Specifically, for Study 2, I recruited psychologists who had experience using the PCL-R in the field to gain a better sense of how well those using the measure could classify PCL-R profiles into their correct empirically supported subtype. Additionally, although participants in Study 1 reported that mean item score profiles were easier to classify, there was not a significant difference in classification accuracy between presentation types. To increase generalizability of the study and to further investigate whether it is useful to have the information in research better align with the information clinicians have in the field, I utilized PCL-R profiles that mirror the score sheet provided in the PCL-R manual (Hare, 2003) for Study 2. These profiles also contained item level data. The methods, results, and discussion of Study 2 are as follows.

CHAPTER III

Study 2

Method

Participants

Participants were psychologists recruited through e-mail requests sent to members of professional organizations and listservs (e.g., Psych and Law Listserv, American Academy of Forensic Psychology), as well as individuals known to conduct forensic evaluations in the field who were thought to be eligible to participate. Participants were also asked to forward the study link to other practicing forensic psychologists who were eligible to participate. Participants who completed the study and provided their email received a $20 Amazon gift card for participation. The final sample of participants consisted of 37 doctoral-level psychologists. Most participants identified as female (n = 27; 72.97%), with some identifying as male (n = 9; 24.32%) and one identifying as non-binary (2.70%). Further, most participants identified as White/European American (n = 28; 75.68%), with three identifying as Black or African American (8.12%), one identifying as Caribbean or Caribbean American (2.70%), two identifying as Hispanic, Latinx, or being from Spanish origin (5.41%), and one identifying as both White/European American and Hispanic, Latinx, or being from Spanish origin (2.70%). Two participants identified as Other (5.41%); one specified "White/UK American" and one did not provide an answer as to how they identified.

Twenty-three participants (62.16%) had obtained a Ph.D., whereas 14 participants (37.84%) had obtained a Psy.D. Thirty-two of the participants reported practicing in the United States (86.49%) and five participants did not identify their country of practice

(13.51%). Most of the participants were licensed (*n* = 36; 97.30%). The one participant who was not licensed reported being trained to administer the PCL-R and had extensive practice using the instrument. Participants had been conducting forensic evaluations for an average of 11.12 years (*SD* = 8.04) and using the PCL-R for an average of 9.09 years (*SD* = 8.23). Participants reported using the PCL-R an average of 7.92 times per year (*SD* = 6.60; three participants who provided a range had their ranges averaged for this calculation).

Most participants reported using the PCL-R for violence risk assessments (*n* = 32; 86.49%) or sexual risk assessments (*n* = 23; 62.16%), but participants also endorsed using the PCL-R for competency to stand trial evaluations (*n* = 2; 5.41%), mental state at the time of the offense evaluations (*n* = 4; 10.81%), mitigation evaluations (*n* = 10; 27.03%), and sexually violent predator evaluations (*n* = 14; 37.84%). Participants were given the option to write in other contexts in which they use the PCL-R, and the following contexts were provided: treatment (*n* = 2; 5.41%), "release decisions" (*n* = 1; 2.70%), and "civil commitment, general mental health in forensic settings, advisability of/amenability to treatment, fitness for duty" (*n* = 1; 2.70%). All participants endorsed partaking in some type of PCL-R training, whether formal or informal. Most of the participants reported having some type of familiarity with PCL-R subtyping (*n* = 23; 62.16%), whereas 14 participants (37.84%) reported they were not previously familiar with the topic. Lastly 11 participants (29.73%) reported that PCL-R subtype findings have impacted how they used or interpreted PCL-R results.

Materials

Materials were presented to participants via Qualtrics, an online survey software platform. A study overview page informed participants that they were completing a study examining the clinical utility of psychopathy subtypes as measured by the PCL-R. They were then asked to read a description of subtyping analyses and the findings of the McCallum et al. (2020) LPA study.

Demographics. Participants completed a demographic survey that contained two sections: 1. General information (e.g., gender, race/ethnicity, highest degree obtained, country of practice) and 2. Career information (e.g., years conducting forensic evaluations, years using the PCL-R, usage of the PCL-R per year, types of forensic evaluations conducted that included the PCL-R, types of PCL-R training attended, familiarity with PCL-R subtyping, impact of PCL-R subtyping on PCL-R interpretation).

PCL-R Profiles. Three profiles for each subtype found by McCallum and colleagues (2020) were presented in Study 2. Specifically, out of the individual profiles presented in Study 1, I randomly selected one profile with a relatively high probability value for the class, one profile with a relatively low probability value for the class, and one profile with a probability value from the middle-third for the class. This process was repeated for each subtype.

Although the participants in Study 1 had a somewhat higher level of classification accuracy when provided profiles of mean item scores than summed facet scores, this difference was not statistically significant. Overall, the participants in Study 1 performed well on both types of presentations of PCL-R profiles. Therefore, the focus of Study 2 was to make the presentation of PCL-R profiles as generalizable to clinical practice as possible. Therefore, the information provided to participants in Study 2 was different than the information provided to participants in Study 1. Study 2 participants were presented with a PCL-R score sheet (Appendix C) created to emulate the score sheet provided in the PCL-R manual. Therefore, for each of the 12 Study 2 profiles, participants were provided the same graphical representation of the prototypical facet scores of each of the four subgroups found by McCallum and colleagues (2020) presented in Study 1 and a PCL-R score sheet with item, facet, and total scores from one offender.

Below each score sheet, participants were asked to classify the profile into one of the four subtype groups. Specifically, they were asked to check the box of the subtype they believe best represented that particular profile. Each participant completed the 12 profile classifications in a randomly assigned order.

Confidence Rating. The PCL-R profile sheet also contained a question asking participants to rate how confident they were in their classification on a scale from 1 to 10 (1 = not confident at all, 10 = extremely confident). Participants were asked to provide a confidence rating after they scored each PCL-R profile.

Open-Ended Questions. At the end of the survey, participants were asked questions to better assess the manner with which they completed the classification task. Specifically, participants were asked to write how they assigned each set of scores to a subtype in as much detail as possible and to express what other information could have been provided to make the task easier. They were also provided a space to communicate to the researchers anything else they wanted to share about the classification task.

Results

Subtype Classification Accuracy

I used descriptive statistics (i.e., percentage values, means) to generally examine classification accuracy. Table 12 provides each participant's classification accuracy across subtypes, including the overall mean and standard deviation for each subtype. Overall, the mean classification accuracy rate across evaluators was 58.56% (*SD* = 15.40). Specifically, accuracy for individual evaluators ranged from 0% to 100% for the prototypic subtype (*M* = 69.37, *SD* = 26.50), from 0% to 100% for the sociopathic subtype (*M* = 45.95, *SD* = 24.03), 0% to 100% for the general offender subtype (*M* = 65.77, *SD* = 40.43), and 0% to 100% for the callous-conning subtype (*M* = 53.15, *SD* = 19.97).

I also used descriptive statistics (i.e., percentage values, means) to examine classification accuracy across subtypes and posterior class probability classification. Table 13 and Figure 3 provide both a qualitative and graphical representation of these findings.

Table 12

Clinician Participants' Classification Accuracy for Each Subtype

Participant	Prototypic	Sociopathic	General	Callous-conning	Overall
1	33.33	33.33	33.33	33.33	33.33
2	100.00	33.33	66.67	66.67	66.66
3	66.67	33.33	100.00	33.33	58.33
4	33.33	66.67	.00	33.33	33.33
5	33.33	.00	100.00	66.67	50.00
6	100.00	33.33	66.67	66.67	66.66
7	100.00	100.00	100.00	66.67	91.66
8	100.00	33.33	.00	66.67	50.00
9	66.67	33.33	100.00	33.33	58.33
10	66.67	.00	66.67	66.67	50.00
11	100.00	66.67	.00	33.33	50.00
12	100.00	33.33	100.00	66.67	75.00
13	66.67	.00	.00	.00	16.66
14	100.00	.00	66.67	100.00	66.66
15	66.67	66.67	66.67	33.33	58.33
16	66.67	66.67	100.00	66.67	75.00
17	100.00	66.67	100.00	33.33	75.00
18	100.00	66.67	33.33	66.67	66.66
19	66.67	66.67	100.00	33.33	66.66
20	100.00	33.33	100.00	66.67	75.00
21	66.67	66.67	100.00	66.67	75.00
22	66.67	66.67	100.00	66.67	75.00
23	66.67	33.33	100.00	33.33	58.33
24	66.67	33.33	100.00	66.67	66.66
25	66.67	66.67	.00	66.67	50.00
26	100.00	33.33	66.67	33.33	58.33
27	66.67	33.33	100.00	66.67	66.66
28	66.67	66.67	66.67	66.67	66.66
29	100.00	33.33	100.00	66.67	75.00
30	33.33	66.67	.00	33.33	33.33
31	33.33	66.67	100.00	33.33	58.33
32	66.67	33.33	66.67	66.67	58.33
33	.00	66.67	.00	66.67	33.33
34	33.33	66.67	33.33	66.67	50.00
35	33.33	33.33	100.00	66.67	58.33
36	66.67	33.33	100.00	33.33	58.33
37	66.67	66.67	.00	33.33	41.67
Mean (SD)	69.37 (26.50)	45.95 (24.03)	65.77 (40.43)	53.15 (19.97)	58.56 (15.40)

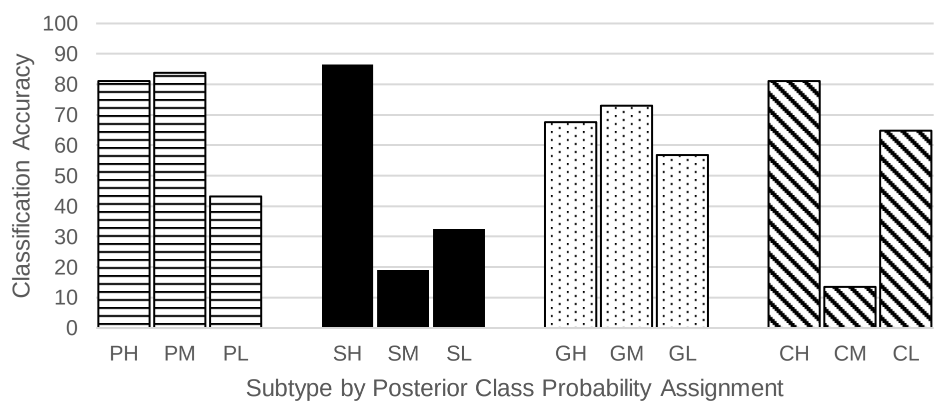
Table 13

Classification Accuracy Across Subtypes and Posterior Class Assignment Categories for Clinician Participants

Posterior Class	Subtype			
	Prototypic	Sociopathic	General	Callous-conning
High	81.1	86.5	67.6	81.1
Moderate	83.8	18.9	73	13.5
Low	43.2	32.4	56.8	64.9
Mean (SD)	69.4 (22.7)	45.9 (35.8)	65.8 (8.2)	53.2 (35.3)

Figure 3

Clinician Participants' Classification Accuracy for each Subtype across each Posterior Class Probability Assignment



Note. Classification across subtypes and posterior class assignment categories. PH = Prototypic subtype, high posterior class assignment; PM = Prototypic subtype, moderate posterior class assignment; PL = Prototypic subtype, low posterior class assignment; SH = Sociopathic subtype, high posterior class assignment; SM = Sociopathic subtype, moderate posterior class assignment; SL = Sociopathic subtype, low posterior class assignment; GH = General Offender subtype, high posterior class assignment; GM = General Offender subtype, moderate posterior class assignment; GL = General Offender subtype, low posterior class assignment. CH = Callous-conning subtype, high posterior class assignment; CM = Callous-conning subtype, moderate posterior class assignment; CL = Callous-conning subtype, low posterior class assignment.

To perform better than chance (i.e., 25%), participants had to correctly classify at least 4 out of 12 profiles (i.e., 25%). Binomial tests (see Table 14) revealed that 25 participants (67.57%) performed better than chance at $p < .05$. Although 12 participants (32.43%) did not perform better than chance at a significant level, six participants (16.22%) had more modest classification accuracies that approached better than chance levels ($p = .054$).

Table 14

Binomial Test p-values for Clinician Participants

Participant	*p*-value
1	.35
2	.003
3	.01
4	0.35
5	.054
6	.003
7	<.001
8	.054
9	.01
10	.054
11	.054
12	<.001
13	.84
14	.003
15	.01
16	<.001
17	<.001
18	.003
19	.003
20	<.001
21	<.001
22	<.001
23	.01
24	.003
25	.054
26	.01
27	.003
28	.003
29	<.001
30	0.35
31	.01
32	.01
33	0.35
34	.054
35	.01

36	.01
37	.16

Posterior Class Assignment Probability Accuracy

As can be seen in Figure 3, participants had differing levels of classification accuracy for the different levels of posterior classification probability. Table 15 provides the descriptive statistics of accuracy across the subtypes and posterior class assignment probability categories when the classification was coded 0 for incorrect and 1 for correct.

Table 15

Mean Accuracy for each Profile when Coded by 0s and 1s for Clinician Participants

Class/Probability Class	Incorrect	Correct	Mean (SD)
Pro High	7	30	.8108 (.39706)
Pro Med	6	31	.8378 (.37368)
Pro Low	21	16	.4324 (.50225)
Soc High	5	32	.8649 (.34658)
Soc Med	30	7	.1892 (.39706)
Soc Low	25	12	.3243 (.47458)
Gen High	12	25	.6757 (.47458)
Gen Med	10	27	.7297 (.45023)
Gen Low	16	21	.5676 (.50225)
Cal High	7	30	.8108 (.39706)
Cal Med	32	5	.1351 (.34658)
Cal Low	13	24	.6486 (.48398)

Note. Pro High = Prototypic offender subtype with high posterior class assignment probability; Pro Med = Prototypic offender subtype with medium posterior class assignment probability; Pro Low = Prototypic offender subtype with low posterior class assignment probability; Soc High = Sociopathic offender subtype with high posterior class assignment probability; Soc Med = Sociopathic offender subtype with medium posterior class assignment probability; Soc Low = Sociopathic offender subtype with low posterior class assignment probability; Gen High = General offender subtype with high posterior class assignment probability; Gen Med = General offender subtype with medium posterior class assignment probability; Gen Low = General offender subtype with low posterior class assignment probability; Cal High = Callous-conning offender subtype with high posterior class assignment probability; Cal Med = Callous-conning offender subtype with medium posterior class assignment probability; Cal Low = Callous-conning offender subtype with low posterior class assignment probability

I used Cochran's Q to determine if there was a statistically significant difference in classification accuracy between the posterior class assignment probabilities within each subtype. If there was a significant difference, I used McNemar's tests to more specifically determine which posterior class probability classifications were significantly different from one another within each subtype. Table 16 provides the results from the Cochran's Q and McNemar's tests used to investigate the differences in classification accuracy between the posterior class assignment probabilities within each subtype.

Table 16

Cochran's Q & McNemar's Statistics for the Comparison of Accuracy Between Posterior Class Assignment Within Each Subtype for Clinician Participants

Class/Probability Class	Cochran's Q	df	*p*-value	OR [95% CI]
Prototypic	17.58	2	<.001	
High & Medium			1.0	1.2 [0.37, 3.93]
High & Low			<.001	15 [1.98, 113.56]
Medium & Low			<.001	6 [1.77, 20.37]
Sociopathic	32.81	2	<.001	
High & Medium			<.001	26 [3.53, 191.60]
High & Low			<.001	21 [2.82, 156.12]
Medium & Low			.302	2 [.68, 5.85]
General	5.09	2	.078	
Cal-Con	29.20	2	<.001	
High & Medium			<.001	-
High & Low			.238	2 [.75, 5.33]
Medium & Low			<.001	5.75 [1.99, 16.63]

Note. Cells in the OR [95% CI] column that have a dash could not be calculated due to a cell value of 0.

For example, I ran Cochran's Q to compare accuracy rates among the three posterior class assignment probabilities within the prototypic subtype. There was a significant difference found among the prototypic class, Q(2) = 17.58, $p < .001$. I then ran

McNemar's tests, which showed that participants had a higher classification accuracy when the profile had a high posterior class assignment compared to a low posterior class assignment ($p < .001$), as well as when provided a profile with a medium posterior class assignment compared to and low posterior class assignment ($p < .001$).

I also ran Cochran's Q to compare accuracy rates among the three posterior class probability assignments within the sociopathic subtype. There was a significant difference found among the sociopathic class, $Q(2) = 32.81$, $p < .001$. I then ran McNemar's tests, which showed that participants had a higher classification accuracy when the profile had a high posterior class assignment compared to a medium posterior class assignment ($p < .001$), as well as compared to when the profile had a low posterior class assignment ($p < .001$).

Additionally, I ran Cochran's Q to compare accuracy rates among the three posterior class probability assignments within the callous-conning subtype. There was a significant difference found among the callous-conning class, $Q(2) = 29.2$, $p < .001$. I then ran McNemar's tests, which showed that participants had a higher classification accuracy when the profile had a high posterior class assignment compared to a medium posterior class assignment ($p < .001$), as well as when provided a profile with a medium posterior class assignment compared to and low posterior class assignment ($p < .001$).

I ran Cochran's Q to compare accuracy rates among the three posterior class probability assignments within the general offender class and did not find a statistically significant difference, $Q(2) = 5.09$, $p = .078$.

Did Classification Accuracy for each Subtype Differ within each Posterior Class Probability Classification?

I also used Cochran's Q to determine if there was a statistically significant difference between subtypes within each posterior class probability assignment. If there was a significant difference, I used McNemar's tests to more specifically determine which subtypes were significantly different from one another within each posterior class probability assignment. Table 17 provides the results of these analyses.

Table 17

Cochran's Q & McNemars Statistics for the Comparison of Accuracy between each Subtype within each Posterior Class Assignment Category for Clinician Participants

Class/Probability Class	Cochran's Q	df	*p*-value	OR [95% CI]
High	4.94	3	.176	
Medium	56.74	3	<.001	
Pro & Soc			<.001	25 [3.39, 184.51]
Pro & Gen			.388	2 [.60, 6.64]
Pro & Cal			<.001	-
Soc & Gen			<.001	-
Soc & Cal			.727	1.67 [.40, 6.97]
Gen & Cal			<.001	-
Low	8.4	3	.038	
Pro & Soc			.481	1.57 [.61, 4.05]
Pro & Gen			.383	1.63 [.67, 3.92]
Pro & Cal			.115	2.33 [.90, 6.07]
Soc & Gen			.108	2.13 [.92, 4.92]
Soc & Cal			.012	4 [1.34, 11.96]
Gen & Cal			.629	1.43 [.54, 3.75]

Note. Pro = Prototypic offender subtype; Soc = Sociopathic offender subtype; Gen = General offender subtype; Cal = Callous-conning subtype; Cells in the OR [95% CI] column that have a dash could not be calculated due to a cell value of 0.

For example, I ran Cochran's Q to compare accuracy rates among the subtypes within the high posterior class probability assignment. There was not a statistically significant difference Q(3) = 4.94, $p = .176$.

I also ran Cochran's Q to compare accuracy rates among the subtypes within the medium posterior class probability assignment. There was a significant difference found among the medium posterior class probability assignment Q(3) = 56.74, $p < .001$. I then ran McNemar's tests, which showed that there was a significant difference in classification accuracy between the medium prototypic profile and medium sociopathic profile ($p < .001$), with higher classification accuracy for the medium prototypic profile. There was also a statistically significant difference between the medium prototypic profile and the medium callous-conning profile ($p < .001$), again with higher classification accuracy for the medium prototypic profile. Further, there were statistically significant differences ($p < .001$) found between the medium general profile and the medium sociopathic profile, as well as the medium callous-conning profile, with higher classification accuracy for the medium general profile.

Additionally, I conducted Cochran's Q to compare accuracy rates among the subtypes within the low posterior class probability assignment. There was a significant difference found among the low posterior class probability assignment Q(3) = 8.41 $p = .038$. When using McNemar's tests, the only statistical difference between the subtypes within the low posterior class assignment was between the sociopathic subtype and the callous-conning subtype ($p = .012$) in which participants we able to more accurately identify low-probability callous-conning profiles than low-probability profiles from the sociopathic class.

Confidence and Accuracy

Participants rated their confidence for each classification on a scale from one to 10 (1 = not confident at all, 10 = extremely confident). I also used descriptive statistics (i.e., percentage values, means) to examine confidence across subtype and posterior class assignment probability. Table 18 provides a qualitative representation of these findings. Figure 4 provides a graphical representation of the mean confidence rating among participants across each subtype and posterior class assignment probability.

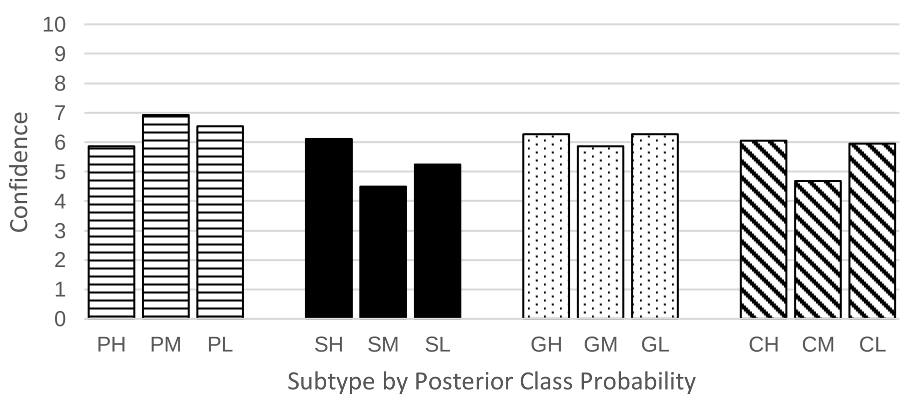
Table 18

Confidence Across Subtypes and Posterior Class Assignment Categories for Clinician Participants

Posterior Class	Subtype			
	Prototypic	Sociopathic	General	Callous-conning
High	5.86	6.12	6.27	6.05
Moderate	6.92	4.49	5.86	4.68
Low	6.54	5.24	6.27	5.95
Mean	6.44	5.28	6.13	5.56
(SD)	(.54)	(.82)	(.24)	(.76)

Figure 4

Clinician Participants' Confidence Rating for each Subtype Across each Posterior Class Probability Assignment



Note. Confidence across subtypes and posterior class assignment categories. PH = Prototypic subtype, high posterior class assignment; PM = Prototypic subtype, moderate posterior class assignment; PL = Prototypic subtype, low posterior class assignment; SH = Sociopathic subtype, high posterior class assignment; SM = Sociopathic subtype, moderate posterior class assignment;

SL = Sociopathic subtype, low posterior class assignment; GH = General Offender subtype, high posterior class assignment; GM = General Offender subtype, moderate posterior class assignment; GL = General Offender subtype, low posterior class assignment. CH = Callous-conning subtype, high posterior class assignment; CM = Callous-conning subtype, moderate posterior class assignment; CL = Callous-conning subtype, low posterior class assignment.

I used correlations to examine the relation between confidence and classification accuracy. Table 19 provides the correlation and *p*-value for that analysis. Confidence for the high prototypic profile was statistically significantly correlated with classification accuracy in that participants were more confident classifying when they were more accurate for the high prototypic profile. Unexpectedly, confidence for the low sociopathic profile was negatively statistically significantly correlated with classification accuracy, in that participants were less confident when they were more accurate for the low sociopathic profile.

I also used a 3 (posterior class probability assignment) x 4 (subtype) two-way MANOVA to further investigate the differences in confidence between posterior class assignment probability and subtype. There was a main effect for subtype $F(3, 34) = 16.40$, $p < .001$, $\eta p^2 = .59$, a main effect for posterior class assignment $F(3, 35) = 11.67$, $p < .001$, $\eta p^2 = .40$, and a two-way interaction between subtype and posterior class assignment, $F(6, 31) = 7.03$, $p < .001$, $\eta p^2 = .58$.

As can be seen in Figure 4, the pattern of associations between confidence and profile posterior class assignment varied across subtypes. Whereas we would expect to see the highest confidence for profiles with high posterior class assignment probabilities and lowest confidence for profiles with low posterior class assignment probabilities (with medium profiles somewhere in between), this pattern was not observed. Specifically, the largest difference in confidence between the posterior class assignment probabilities within the prototypic subtype was between the medium (*M* = 6.92, *SD* = 1.72) and high (*M* = 5.86, *SD* = 2.16) posterior class assignments (Cohen's *d* = .54).

Table 19

Correlation Between Accuracy and Confidence for Clinician Participants

Posterior Class	r	*p*-value
Pro High	.325	.049
Pro Med	.238	.156
Pro Low	-.081	.633
Soc High	.164	.322
Soc Med	.236	.159
Soc Low	-.488	.005
Gen High	.292	.079
Gen Med	.131	.439
Gen Low	.220	.190
Cal High	.237	.102
Cal Med	-.267	.111
Cal Low	.201	.232

Note. Pro High = Prototypic offender subtype with high posterior class assignment probability; Pro Med = Prototypic offender subtype with medium posterior class assignment probability; Pro Low = Prototypic offender subtype with low posterior class assignment probability; Soc High = Sociopathic offender subtype with high posterior class assignment probability; Soc Med = Sociopathic offender subtype with medium posterior class assignment probability; Soc Low = Sociopathic offender subtype with low posterior class assignment probability; Gen High = General offender subtype with high posterior class assignment probability; Gen Med = General offender subtype with medium posterior class assignment probability; Gen Low = General offender subtype with low posterior class assignment probability; Cal High = Callous-conning offender subtype with high posterior class assignment probability; Cal Med = Callous-conning offender subtype with medium posterior class assignment probability; Cal Low = Callous-conning offender subtype with low posterior class assignment probability

In contrast, the largest difference between the posterior class assignment probabilities within the sociopathic subtype was between the high (*M* = 6.12, *SD* = 1.73) and medium (*M* = 4.49, *SD* = 1.95) posterior class probability assignments (Cohen's *d* = .88), and the largest difference between the posterior class assignment probabilities within the callous-conning subtype was between the high (*M* = 6.05, *SD* = 1.63) and medium (*M* = 4.68, *SD* = 1.62) posterior class probability assignments (Cohen's *d* = .84).

As can be seen in Figure 4, the pattern of average confidence across the subtypes and posterior class assignments looks similar to the pattern depicted in Figure 3 of accuracy across the subtypes and posterior class assignment probabilities. For example, the levels of confidence displayed in Figure 4 for the sociopathic and callous-conning subtypes match the pattern for the levels of accuracy for those subtypes displayed in Figure 3. More specifically, for both subtypes, the bar for the high posterior class assignment is higher than the low posterior class assignment, which is higher than the medium posterior class assignment. In contrast, however, the pattern of confidence across posterior class assignment probabilities within the prototypic and general offender subtypes depicted in Figure 4 does not match the pattern of accuracy across posterior class assignment probabilities within these subtypes.

Summary of Open-Ended Responses

Participants were asked to describe how they approached the classification task. As with Study 1, answers varied. Many participants mapped the scores they were given onto the graphs provided to make their classification. Some participants mentioned taking total score into consideration and others either calculated their own math or drew their own graphs to better address the task. Further, other participants wrote that they had inconsistent strategies throughout the classification task, meaning that they based their strategy on the profile they were presented. When asked if there was any other information that they would have liked to have to complete the classification task, participants reported that they would have liked to have more clinical information (e.g., information collected from a clinical interview), plotted graphs with t-scores, more information about the subtypes found in the study conducted by McCallum et el. (2020), more clear directions on how to utilize subscale scores, and a tool to graph each individual profile.

Profile Comparisons Between Studies

Because the 12 profiles rated by the clinician participants in Study 2 were among the 120 profiles rated by the doctoral student participants in Study 1, I was able to compare accuracy rates across the two studies for this subsample of profiles. As noted in Table 20, there were some profiles for which the classification rate was higher for the student participants than the clinician participants. Both students and clinicians were able to classify high probability profiles with a high degree of accuracy across all subtypes; however, differences emerged between the groups of participants for profiles that were less prototypic. For example, for the student participants, the sociopathic profile with the moderate posterior class assignment obtained a classification rate of 50% and 75% for the mean item and summed facet presentations respectively; however, the clinician accuracy rate was only 18.9%. However, there were also some profiles for which the classification rate was higher for the clinician participants than the student participants. For example, for the prototypic profile with the low posterior class assignment, clinicians earned a classification rate of 43.2%, whereas the mean item and summed facet classification rates were both 16.7% for the student participants.

Table 20

Percent Correct for Each Type of Presentation Method for Each of the 12 Profiles Presented in Both Study 1 and Study 2

	ProH	ProM	ProL	SocH	SocM	SocL	GenH	GenM	GenL	CalH	CalM	CalL
Student Mean Item	100	66.7	16.7	100	50	41.7	91.7	75	67	100	75	91.7
Student Summed Facet	91.7	50	16.7	100	75	41.7	83.3	83.3	41.7	100	41.7	91.7
Clinicians	81.1	83.8	43.2	86.5	18.9	32.4	67.6	73	56.8	81.1	13.5	64.9

Note. All numbers are percentage values. ProH = Prototypic offender subtype with high posterior class assignment probability; ProM = Prototypic offender subtype with medium posterior class assignment probability; ProL = Prototypic offender subtype with low posterior class assignment probability; SocH = Sociopathic offender subtype with high posterior class assignment probability; SocM = Sociopathic offender subtype with medium posterior class assignment probability; SocL = Sociopathic offender subtype with low posterior class assignment probability; GenH = General offender subtype with high posterior class assignment probability; GenM = General offender subtype with medium posterior class assignment probability; GenL = General offender subtype with low posterior class assignment probability; CalH = Callous-conning offender subtype with high posterior class assignment probability; CalM = Callous-conning offender subtype with medium posterior class assignment probability; CalL = Callous-conning offender subtype with low posterior class assignment probability

Discussion

The second part of this study asked clinicians to classify PCL-R profiles using score sheets similar to those used in real world settings in order to generalize preliminary findings observed in Study 1. Clinicians performed moderately well overall; however, they performed worse than hypothesized. Contrary to what was expected and what was observed among the doctoral student participants, forensic evaluator participants were better able to classify profiles from the prototypic subtype compared to the other subtypes. Participants had the most difficulty classifying profiles from the sociopathic and callous-conning subtypes. It is possible that the sociopathic and callous-conning subtypes were the most difficult to classify because the total scores of these profiles are the most variable, yet also the closest to the middle.

Generally, among the prototypic, sociopathic and callous-conning subtypes, classification accuracy was greater for profiles that had a higher posterior class assignment probability. For example, a majority of participants were able to accurately classify the high posterior class probability sociopathic (n = 32; 86.49%) and callous-conning (n = 30; 81.08%) profiles. In contrast, only a small minority of participants were able to accurately classify the moderate posterior class probability sociopathic (n = 7; 18.9%) and callous-conning (n = 5; 13.5%) profiles. These findings continue to raise concern for the accuracy with which clinicians can classify profiles that are not as prototypical within their subtype and suggests further research is needed to investigate which elements of particular subtypes assist in mitigating classification difficulties.

Additionally, unlike the doctoral student participants, forensic evaluator participants were generally unable to accurately appraise their ability to perform the task, as confidence ratings and classification accuracy were generally uncorrelated, with the exception of the high prototypic profile in which participants were more confident in their accurate classifications compared to their inaccurate classifications. In fact, for the low sociopathic profile, greater confidence was associated with *lower* classification accuracy. Further, participants were overall more confident with their classifications for profiles with higher posterior class probabilities in all subtypes but the general subtype. It is interesting that confidence among the general offender subtype was not significantly related with accuracy or posterior class probability yet was the second highest correctly classified subtype and the subtype for which I hypothesized participants would perform the best.

Overall, clinicians did not report a uniform method for classification. Interestingly, the clinicians received more detailed information in the profiles presented for classification compared to the student participants (e.g., item level data); however, some participants still reported that having more information may have allowed for easier classification. The other responses from participants regarding other information that they would have wanted to better classify the profiles suggest that training regarding psychopathy and PCL-R subtyping may benefit clinicians in their understanding and accuracy of subtype classification.

CHAPTER IV

General Discussion

These studies were a first step in investigating the clinical utility of psychopathy subtypes using the PCL-R. First, investigators asked clinical psychology doctoral students to classify PCL-R profiles into their correct subtype when provided one of two presentation styles: mean item scores or summed facet scores. Then a follow-up study was conducted in which forensic evaluators were asked to classify a smaller set of profiles using a more ecologically valid manner of information presentation. Overall, findings from these studies suggest it may be possible for psychopathy subtypes to be correctly identified for use in applied clinical contexts; however, this may only be true for individuals with PCL-R profiles that are highly prototypical of empirically derived subtypes. Further, there was a high degree of variability in individual ability to accurately classify profiles and substantial rates of misclassification, especially among the forensic evaluators. Thus, additional research is necessary before psychopathy subtypes can confidently be applied in high-stakes real-world settings.

Although graduate student participants did well in their classifications (i.e., chance would have led to 25% accurate, 75% inaccurate), there was room for improvement with about 30% of profiles being inaccurately classified. Overall, students were better at classifying profiles presented as mean item scores compared to summed facet scores. Although not a significant difference, this supports the hypothesis that it may be easier to classify the profiles when they are presented in a manner that is commensurate with how subtype findings are presented in the research. This idea was reinforced by some participants reporting that they found it easier to classify the mean item scores. This mismatch in presentation of information between research and practice was further investigated in Study 2, as clinicians were asked to complete a smaller classification task using PCL-R profile sheets based on how the PCL-R would be scored in real world settings. Compared to the student participants, the clinicians did not perform as well overall.

Accuracy was variable across participants in both studies. Graduate student participants were overall more accurate when classifying the general offender and callous-conning profiles when presented with mean item scores and sociopathic profiles when presented with summed facet scores. Graduate student participants appeared to perform the worst when provided profiles from the prototypic subtype. While some profiles in the prototypic class will have fairly uniform elevations across facets, others will differ somewhat across facets (i.e., relatively higher on facets 1 and 2 or relatively higher on facets 3 and 4) which might resemble callous-conning or sociopathic subtypes, respectively. Such configurations of facet elevations among individuals with high overall PCL-R scores are consistent with primary and secondary subtypes of psychopathy, and it is likely that the prototypic class from the current project subsumes these variants. Some may even have facet elevations that don’t clearly follow any of the subtypes (e.g., relatively higher on facets 1 and 4). Given this heterogeneity, any time the facet elevations are not uniformly high for an individual within the prototypic subtype, the likelihood of misclassification increases. Despite these potentially nuanced difficulties within the prototypic class, the students’ low classification accuracy among this subtype is concerning due to the external correlates associated with the prototypic subtype. More research is needed to examine the difficulties presented by prototypic profiles in the context of subtype classification.

Interestingly, however, clinician participants did not appear to be as affected by the nuanced differences of the prototypic subtype like the student participants; the clinicians performed the best with the prototypic subtype profiles followed by the general offender subtype. It may be that clinicians are utilizing the PCL-R to make a broader determination of whether someone is or is not high in psychopathic traits, which would make it easier to identify the prototypic and general offender subtypes. By contrast, clinician participant accuracy for the callous-conning and sociopathic subtype profiles was significantly lower compared to the prototypic subtype profiles.

When investigating the misclassifications, callous-conning profiles were generally more likely to be misclassified as prototypic, whereas sociopathic profiles were generally more likely to be misclassified as general offender profiles. This finding is interesting in the context of the debate regarding the centrality of Factor 2 to the construct of psychopathy, and seems to support the idea that, in practice, the more behavioral components are seen to be less central to the construct (Skeem & Cook, 2010). This finding is also cause for concern. Specifically, research has shown that prototypic offenders score higher on violence risk measures compared to the other subtypes (Lehmann et al., 2019; McCallum et al., 2020; Neumann et al., 2016), with sociopathic offenders showing the second highest level of violence risk (Krstic et al., 2018; Lehmann et al., 2019; McCallum et al., 2020). Therefore, if an evaluator misclassifies a callous-conning profile as prototypic, they are overpredicting risk for that individual. In contrast, if an evaluator misclassifies a sociopathic profile as general offender, they would be underpredicting risk for that individual. These misclassifications could have serious consequences for the individuals being evaluated and the communities in which they live. It may be that there are misperceptions on what aspects of psychopathy as measured by the PCL-R are the most relevant, with too much emphasis being placed onto Factor 1 when it is Factor 2 that has a moderate association with antisocial conduct (Leistico et al., 2008), institutional misconduct (Guy et al., 2005), general and violent recidivism (Walters, 2003), and reactive violence (Blais et al., 2014).

One finding consistent across both the student and clinician participants is that posterior class assignment probability of the profiles was influential in the participants' ability to correctly classify the profile. For student participants, the higher the posterior class assignment probability of the profile, the more accurately student participants were able to place it into the correct subtype. For the practitioner participants, the conclusion was not as obvious, as there was surprising variability. For example, practitioners did a significantly better job classifying the callous-conning profile from the low posterior class assignment probability when compared to the profile from the medium posterior class assignment probability. However, across all subtypes, clinicians generally performed well when the posterior class assignment was high (over 67% accuracy).

It is reassuring to find that clinicians are good at identifying profiles that look like the class to which they have been assigned; however, it casts some doubt on the ability for clinician's to accurately identify the subtype of profiles that may look slightly different than the "typical" profile for that class. More specifically, despite there being some profiles that are highly typical of their respective classes for which both students and practitioners were able to correctly classify, there are also cases within each class that are less typical that were more difficult to classify. Researchers can classify these nontypical profiles using large samples and complex statistical models (e.g., LPA, cluster analysis) that allow for the simultaneous comparison of the probability of belonging to each possible class, ultimately selecting the class that is empirically the best fit. Although the posterior class assignment for one class may not be much higher than it is for another class (e.g., .54 for callous-conning and .48 for prototypic), these analyses still place the profile into a class. However, evaluators do not have access to these large data sets and are not typically trained to apply these latent variable models; they are making subtype determinations based on the single case in front of them, making the best classification based on the information in that single case. Therefore, unless it is a highly typical profile, it may not be realistic to expect clinicians to use subtypes as their mode of profile interpretation; more traditional PCL-R interpretations may be warranted. Further research should continue to investigate elements that lead to misclassification in profiles with lower prototypicality.

Additionally, literature has posited that clinicians often have a blind spot when it comes to how they believe they are performing, in that many clinicians are overconfident with their clinical judgment (Borum et al., 1993). Although student participants had more confidence when their classification was accurate than when it was inaccurate, for the practitioner participants, confidence and accuracy were generally uncorrelated. This suggests that practitioners using the PCL-R may be over- or underconfident about their conclusions for some types of profiles (e.g., callous-conning profiles, general offender profiles). This could potentially lead to misinterpretations of PCL-R data and, particularly in high-stakes forensic settings, presentation of misrepresented data to factfinders who opine on important forensic decisions (e.g., sentencing).

As aforementioned, student participants were overall better able to classify the profiles they were given compared to the clinician participants. It is possible that study design may have inherently influenced this finding. More specifically, the profiles presented to students contained less information than the profiles presented to the clinicians, as the profiles presented to the clinicians included item level data. Therefore, it could be that more information creates more extraneous information to sift through leading the clinician to approach the classification task for each profile in a slightly different manner based on this extraneous information. Additionally, if that is true it would be a juxtaposition to the clinician participants' request for more information to better inform their classification. Further, the classification task is purely based on numerical values; therefore, although the clinician participants requested more information to assist in making their classifications, more information should be unnecessary.

Further, the student participants were provided a training before undergoing the classification task. It could be that student participants learned from the training, whereas the short overview provided to the clinicians did not provide enough education prior to the classification task. However, in order to enhance the ecological validity of the study, I did not provide extensive training to the clinician participants prior to the classification task, as most practitioners do not receive specific training regarding psychopathy subtyping prior to performing their clinical work. Lastly, student participants were provided 10 times the number of profiles to classify. It is possible that student participants learned from their own classification throughout the task. More research is needed to identify the elements that help or hinder classification of PCL-R profiles into their correct subtype.

Limitations

The biggest limitation to the current studies was the small sample size. However, this study serves as a strong foundation for further research to continue to investigate the clinical utility of PCL-R subtypes. Another limitation is that, although the researchers attempted to make the materials in the second study generalizable by using a PCL-R score sheet similar to the score sheet from the PCL-R manual (Hare, 2003), many of the clinicians noted in their responses that they would have liked to have more contextual information for the individual represented by the profile. Future research should continue with study designs that mirror more closely how clinicians are practicing in the field to further increase ecological validity. Lastly, the profiles used for this study come from SVP evaluations. It is possible that there is something different about these evaluations or individuals being evaluated compared to a more generalized sample that could have affected our participants ability to classify these profiles. Future studies should utilize PCL-R profiles from a more general subset of evaluations.

CHAPTER V

Conclusion

In sum, this study aimed to investigate the clinical utility of psychopathy subtypes, particularly given the use of the PCL-R and its subtypes in high stakes forensic settings such as capital cases (Olver et al., 2020). Overall, findings show that although individuals may be able to correctly classify PCL-R profiles into their correct subtype, this ability is dependent on many factors, such as the subtype to which the profile belongs or the prototypicality of the profile. Further, participants reported using various strategies in their classification task. Although it is positive to see that the participants were thinking critically, the use of such a wide range of strategies may have impacted the variability in participants' level of accuracy. Better training and education on best practices in regard to classification and interpretation of subtyping findings may benefit practitioners in the field.

REFERENCES

Blais, J., Solodukhin, E., & Forth, A. E. (2014). A meta-analysis exploring the relationship between psychopathy and instrumental versus reactive violence. *Criminal Justice and Behavior, 41*(7), 797–821.

Boccaccini, M. T., Chevalier, C., Murrie, D. C., & Varela, J. G. (2017). Psychopathy Checklist-Revised use and reporting practices in sexually violent predator evaluations. *Sexual Abuse, 29*(6), 592-614.

Boduszek, D., Debowska, A., Dhingra, K., & DeLisi, M. (2016). Introduction and validation of Psychopathic Personality Traits Scale (PPTS) in a large prison sample. *Journal of Criminal Justice, 46,* 9–17. https://doi.org/10.1016/j.jcrimjus.2016.02.004

Boduszek, D., Debowska, A., & Willmott, D. (2017). Latent profile analysis of psychopathic traits among homicide, general violent, property, and white-collar offenders. *Journal of Criminal Justice, 51,* 17–23. https://doi.org/10.1016/j.jcrimjus.2017.06.001

Borum, R., Otto, R., & Golding, S. (1993). Improving clinical judgment and decision making in forensic evaluation. *Journal of Psychiatry and Law, 21*(1), 35–76.

Cleckley, H. (1941). *The mask of sanity*. Mosby.

DeMatteo, D., Hart, S. D., Heilbrun, K., Boccaccini, M. T., Cunningham, M. D., Douglas, K. S., Dvoskin, J. A., Edens, J. F., Guy, L. S., Murrie, D. C., Otto, R. K., Packer, I. K., & Reidy, T. J. (2020). Statement of concerned experts on the use of the Hare Psychopathy Checklist-Revised in capital sentencing to assess risk for institutional violence. *Psychology, Public Policy, and Law*, *26*(2), 133–144.

Driessen, J. M. A., Fanti, K. A., Glennon, J. C., Neumann, C. S., Baskin-Sommers, A. R., & Brazil, I. A. (2018). A comparison of latent profiles in antisocial male offenders. *Journal of Criminal Justice, 57*, 47–55. https://doi.org/10.1016/j.jcrimjus.2018.04.001

Drislane, L. E., Patrick, C. J., Sourander, A., Sillanmäki, L., Aggen, S. H., Elonheimo, H., & ... Kendler, K. S. (2014). Distinct variants of extreme psychopathic individuals in society at large: Evidence from a population-based sample. *Personality Disorders: Theory, Research, and Treatment*, *5*(2), 154-163. https://doi.org/10.1037/per0000060

Falkenbach, D. M., Reinhard, E. E., & Larson, F. R. R. (2017). Theory based gender differences in psychopathy subtypes. *Personality and Individual Differences*, *105*, 1–6. https://doi.org/10.1016/j.paid.2016.09.023

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

Fraley, C. & Raftery, A. E. (1998) How many clusters? Which clustering method? Answers via model-based cluster analysis. *The Computer Journal, 41*(8), 578–588. https://doi.org/10.1093/comjnl/41.8.578

Guy, L. S., Edens, J. F., Anthony, C., & Douglas, K. S. (2005). Does psychopathy predict institutional misconduct among adults? A meta-analytic investigation. *Journal of Consulting and Clinical Psychology*, *73*(6), 1056–1064. https://doi.org/10.1037/0022-006X.73.6.1056

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

Hare, R. D. (2003). *Hare PCL-R* (2nd edition). Multi-Health Systems.

Hart, S., Cox, D., & Hare, R. D. (1995). *Manual for the Psychopathy Checklist: Screening Version (PCL: SV)*. Multi-Health Systems.

Hare, R. D., & Neumann, C. S. (2008). Psychopathy as a clinical and empirical construct. *Annual Review of Clinical Psychology*, *4*, 217–246. https://doi.org/10.1146/annurev.clinpsy.3.022806.091452

Hare, R. D., & Neumann, C. S. (2010, January 1). The role of antisociality in the psychopathy construct: Comment on Skeem and Cooke (2010). *Psychological Assessment*, *22*(2), 446-454.

Hare, R. D., Neumann, C. S., & Mokros, A. (2018). The PCL-R assessment of psychopathy: Development, structural properties, and new directions. In C. Patrick (Ed.), *Handbook of psychopathy* (2nd ed.). Guilford Press.

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

Hicks, B. M., & Drislane, L. E. (2018). Variants ("Subtypes") of Psychopathy. In C. J. Patrick (Ed.), *Handbook of psychopathy* (2nd ed., pp. 311-337). Guilford Press.

Hicks, B. M., Markon, K. E., Patrick, C. J., Krueger, R. F., & Newman, J. P. (2004, January 1). Identifying psychopathy subtypes on the basis of personality structure. *Psychological Assessment*, *16*(3), 276–288.

KARPMAN, B. (1946). Psychopathy in the scheme of human typology. *Journal of Nervous and Mental Disease, 103,* 276–288. http://doi.org/10.1097/00005053-194603000-00007

Karpman, B. (1948). Conscience in the psychopath: Another version. *American Journal of Orthopsychiatry, 18*(3), 455–491. http://doi.org/10.1111/j.1939-0025.1948.tb05109.x

Klein Haneveld, E., Neumann, C. S., Smid, W., Wever, E., & Kamphuis, J. H. (2018). Treatment responsiveness of replicated psychopathy profiles. *Law and Human Behavior*, *42*(5), 484–495. https://doi.org/10.1037/lhb0000305

Krstic, S., Neumann, C. S., Roy, S., Robertson, C. A., Knight, R. A., & Hare, R. D. (2018). Using latent variable- and person-centered approaches to examine the role of psychopathic traits in sex offenders. *Personality Disorders-Theory Research and Treatment*, *9*(3), 207–216. https://doi.org/10.1037/per0000249

Lehmann, R., Neumann, C. S., Hare, R. D., Biedermann, J., Dahle, K. P., & Mokros, A. (2019). A latent profile analysis of violent offenders based on PCL-R factor scores: Criminogenic needs and recidivism risk. *Frontiers in Psychiatry*, *10*, 627. https://doi.org/10.3389/fpsyt.2019.00627

Leistico, A.-M. 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. https://doi.org/10.1007/s10979-007-9096-6

Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology, 68*(1), 151–158. https://doi.org/10.1037/0022-3514.68.1.151

Lilienfeld, S. O., & Widows, M. R. (2005). *Psychopathic Personality Inventory–Revised (PPI-R): Professional manual.* Psychological Assessment Resources

McCallum K. E., Boccaccini, M. T., Varela, J. G., & Turner, D. B. (2021). Psychopathy profiles and Personality Assessment Inventory Scores in a sex offender risk assessment field setting. *Assessment.* https://doi.org/10.1177/10731911211015312

Morey, L. C. (1991). *Personality Assessment Inventory: Professional manual.* Psychological Assessment Resources.

Mokros, A., Hare, R. D., Neumann, C. S., Santila, P., Habermeyer, E., & Nitschke, J. (2015, January 1). Variants of psychopathy in adult male offenders: A latent profile analysis. *Journal Of Abnormal Psychology*, *124*(2), 372-386.

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.

Neumann, C. S., Vitacco, M. J., & Mokros, A. (2016). Using both Variable-Centered and Person-centered approaches to understanding psychopathic personality. In C. B. Gacono (Ed.), *The clinical and forensic assessment of psychopathy: A practitioner's guide* (2nd ed., pp. 14–31). Routledge.

Olver, M. E., Sewall, L. A., Sarty, G. E., Lewis, K., & Wong, S. C. (2015, January 1). A cluster analytic examination and external validation of psychopathic offender subtypes in a multisite sample of Canadian federal offenders. *Journal of Abnormal Psychology*, *124*(2), 355–371.

Olver, M. E., Stockdale, K., Neumann, C., Hare, R., Mokros, A., Baskin-Sommers, A., Brand, E., Folino, J., Gacono, C., Gray, N., Kiehl, K., Knight, R., Leon-Mayer, E., Long, M., Meloy, J. R., Roy, S., Salekin, R., Snowden, R., Thomson, N.,...Yoon, D. (2020). Reliability and validity of the Psychopathy Checklist-Revised in the assessment of risk for institutional violence: A cautionary note on DeMatteo et al. (2020). *Psychology, Public Policy, and the Law*, *26*(4), 490-510

Olver, M. E., & Wong, S. C. P. (2009, January 1). Therapeutic responses of psychopathic sexual offenders: Treatment attrition, therapeutic change, and long-term recidivism. *Journal Of Consulting and Clinical Psychology*, *77*(2), 328–336.

Patrick, C. J. (2010). Triarchic Psychopathy Measure (TriPM). www.phenxtoolkit.org/index.php?pageLink=browse.protocoldetails&id=121601

Patrick, C. J. (2018). *Handbook of psychopathy* (2nd ed). Guilford Press.

Patrick, C. J., Fowles, D. C., & Krueger, R. F. (2009, January 1). Triarchic conceptualization of psychopathy: Developmental origins of disinhibition, boldness, and meanness. *Development and Psychopathology*, *21*(3), 913–938.

Paulhus, D. L., Neumann, C. S., & Hare, R. D. (2015). *Manual for the Self-Report Psychopathy Scale* (4th ed.). Multi-Health Systems.

Pinel, P. (1962). *A treatise on insanity* (D. Davis, Trans.). Hafner. (Original work published 1806)

Sewall, L. A., & Olver, M. E. (2019). Psychopathy and treatment outcome: Results from a sexual violence reduction program. *Personality Disorders: Theory, Research,* *and Treatment*, *10*(1), 59–69. https://doi.org/10.1037/per0000297

Skeem, J. L., & Cooke, D. J. (2010, January 1). Is criminal behavior a central component of psychopathy? Conceptual directions for resolving the debate. *Psychological Assessment*, *22*(2), 433– 445.

Skeem, J., Johansson, P., Andershed, H., Kerr, M., & Louden, J. E. (2007). Two subtypes of psychopathic violent offenders that parallel primary and secondary variants. *Journal of Abnormal Psychology, 116*(2), 395-409. https://doi.org/10.1037/0021-843X.116.2.395

Swogger, M. T. & Kosson, D. S. (2007). Identifying subtypes of criminal psychopaths: A Replication and extension. *Criminal Justice and Behavior, 34*(8), 953-970.

Vassileva, J., Kosson, D. S., Abramowitz, C., & Conrod, P. (2005). Psychopathy versus psychopathies in classifying criminal offenders. *Legal and Criminological Psychology, 10*(1), 27-44.

Viljoen, J. L., McLachlan, K., & Vincent, G. M. (2010, January 1). Assessing violence risk and psychopathy in juvenile and adult offenders: A survey of clinical practices. *Assessment*, *17*(3), 377-395.

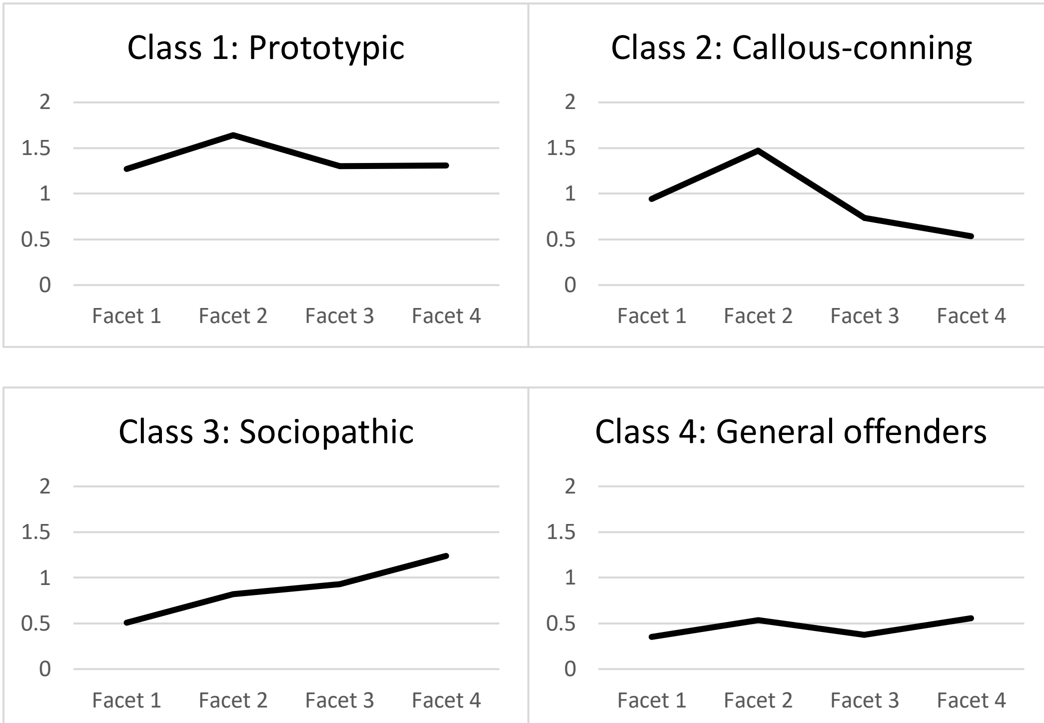
Walters, G. D. (2003b). Predicting institutional adjustment and recidivism with the Psychopathy Checklist Factor Scores: A meta-analysis. *Law and Human Behavior*, *27*(5), 541–558.

Wang, M., Gong, J., Gao, Y., Zhang, X., Yang, W., & Luo, J. (2020). Variants of psychopathy in Chinese male offenders: A latent profile analysis in a large prison sample. *Journal of Criminal Justice, 69*. https://doi.org/10.1016/j.jcrimjus.2020.101708

Williams, G. A., & Kibowski, F. (2016). Latent class analysis and latent profile analysis. In L. A. Jason & D. S. Glenwick (Eds.), *Handbook of methodological approaches to community-based research: Qualitative, quantitative, and mixed methods* (pp. 143–151). Oxford University Press

APPENDIX A

Example PCL-R Plots Using Mean Item Scores



Based on the above classes, please classify the following profile by checking the appropriate box below:

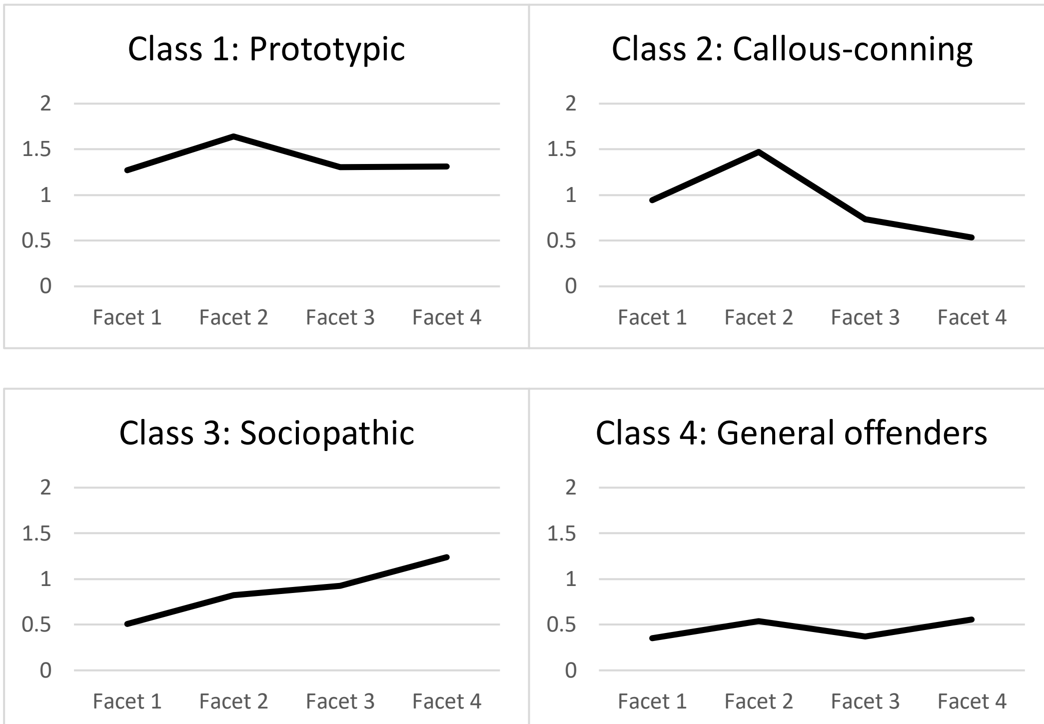
Total Score	8
Facet 1	0
Facet 2	.75
Facet 3	0
Facet 4	.2

☐ Class 1: Prototypic
☐ Class 2: Callous Conning
☐ Class 3: Sociopathic
☐ Class 4: General

On a scale from 1-10, how confident are you in your classification of the above profile (1 = not confident at all, 10 = extremely confident)?

APPENDIX B

<u>Example PCL-R Plots Using Summed Facet Scores</u>



Based on the above classes, please classify the following profile by checking the appropriate box below:

Total Score	8
Facet 1	0
Facet 2	3
Facet 3	0
Facet 4	1

☐ Class 1: Prototypic
☐ Class 2: Callous Conning
☐ Class 3: Sociopathic
☐ Class 4: General

On a scale from 1-10, how confident are you in your classification of the above profile (1 = not confident at all, 10 = extremely confident)?

APPENDIX C

PCL-R Score Sheet
(USE ONLY FOR RESEARCH PURPOSES)

	Facet 1	Facet 2	Facet 3	Facet 4	Total Score
1. Glibness/Superficial Charm	X				X
2. Grandiose Sense of Self-Worth	X				X
3. Need for Stimulation/Boredom			X		X
4. Pathological Lying	X				X
5. Conning/Manipulative	X				X
6. Lack of Remorse or Guilt		X			X
7. Shallow Affect		X			X
8. Callous/Lack of Empathy		X			X
9. Parasitic Lifestyle			X		X
10. Poor Behavioral Controls				X	X
11. Promiscuous Sexual Behavior					X
12. Early Behavioral Problems				X	X
13. Lack of Realistic Long-term Goal			X		X
14. Impulsivity			X		X
15. Irresponsibility			X		X
16. Failure to Accept Responsibility		X			X
17. Many Marital Relationships					X
18. Juvenile Delinquency				X	X
19. Revocation Conditional Release				X	X
20. Criminal Versatility				X	X

	Facet 1	Facet 2	Facet 3	Facet 4	Sum Facets 1&2	Sum Facets 3&4	Total Score
Totals	X	X	X	X	X	X	X
Number of Omitted Items	X	X	X	X	X	X	X
Adjusted Scores	X	X	X	X	X	X	X

APPENDIX D

Sam Houston
STATE UNIVERSITY
A Member of The Texas State University System

Date: Jul 17, 2020 3:34 PM CDT

TO: Gabrielle Trupp Marcus Boccaccini
FROM: SHSU IRB
PROJECT TITLE: Clinical Utility of Psychopathy Subtypes Based on Latent Profile Analysis
PROTOCOL #: IRB-2020-137
SUBMISSION TYPE: Initial
ACTION: Approved
DECISION DATE: July 10, 2020
ADMINISTRATIVE CHECK-IN DATE: July 10, 2021
EXPEDITED REVIEW CATEGORY: 7. Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

SPECIAL IRB UPDATE REGARDING THE COVID-19 CRISIS: Although this study is approved, please note that face-to-face human subject research must be paused until the CDC and SHSU has determined that the current COVID-19 crisis has passed. This pause is effective immediately. Approved online human subject research may continue. If you have an approved face-to-face study and deem it feasible to move the study to online data collection, please submit a Modification through Cayuse. Indicate in the Modification that the change is being implemented as a COVID-19 safety precaution to help the IRB prioritize the submission. The IRB will continue reviewing applications unless we are advised to do otherwise.

Greetings,

The above-referenced submission has been reviewed by the IRB and it has been Approved. Because this study received expedited review and the IRB determined that a renewal submission is not needed, this decision does not necessarily expire; however, you will be receiving an email notification on the anniversary of this study approval, which will be on July 10, 2021 (NOTE: please review the reminder information below regarding Study Administrative Check-In). This study approval is based on an appropriate risk/benefit ratio and a project design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

Since Cayuse IRB does not currently possess the ability to provide a "stamp of approval" on any recruitment or consent documentation, it is the strong recommendation of this office to please include the following approval language in the footer of those recruitment and consent documents: IRB-2020-137/July 10, 2020/July 10, 2021.

Please remember that informed consent is a process beginning with a description of the project and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the project via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Modifications: Please note that any revision to previously approved materials must be approved by this committee prior to initiation. Please submit a Modification Submission through Cayuse IRB for this procedure.

Incidents: All UNANTICIPATED PROBLEMS involving risks to subjects or others and SERIOUS and UNEXPECTED adverse events must be reported promptly to this office. Please submit an Incident Submission through Cayuse IRB for this procedure. All Department of Health and Human Services and sponsor reporting requirements should also be followed.

Study Administrative Check-In: Based on the risks, this project does not require renewal. Rather, you are required to administratively check in with the IRB on an annual basis. July 10, 2021 is the anniversary of the review of your protocol. The following are the conditions of the IRB approval for IRB-2020-137 Clinical Utility of Psychopathy Subtypes Based on Latent Profile Analysis .

1. When this project is finished or terminated, a Closure submission is required.

2. Changes to the approved protocol require prior board approval (NOTE: see the directive above related to Modifications).

3. Human subjects training is required to be kept current at citiprogram.org by renewing training every 5 years.

4. If incidents (i.e., adverse events) or unanticipated problems involving risks to subjects or others (UPIRSO) (e.g., data collected unintentionally without obtaining informed consent) have occurred during this approval period, you are required to submit a Incident to report the adverse event or UPIRSO to the IRB.

Please note that all research records should be retained for a minimum of three years after the completion of the project. If you have any questions, please contact the Sharla Miles at 936-294-4875 or irb@shsu.edu. Please include your protocol number in all correspondence with this committee.

Sincerely,

Chase Young, Ph.D.
Chair, IRB
Hannah R. Gerber, Ph.D.
Co-Chair, IRB

VITA

Gabriele F. Trupp

EDUCATION

Sam Houston State University, Huntsville, TX

Doctor of Philosophy (Clinical Psychology), June 2022 (Anticipated)
Dissertation: *Clinical Utility of Psychopathy Subtypes Based on Latent Profile Analysis*

CUNY John Jay College of Criminal Justice, New York, NY

Master of Arts (Forensic Psychology), June 2017
Thesis: *Treatment, Supervision, and Recidivism of Individuals Convicted of a Sex Offense in the United States: A Pilot Study*

University of Michigan, Ann Arbor, MI

Bachelor of Arts (Major: Psychology; Minor Crime and Justice), May 2014

DOCTORAL CLINICAL EXPERIENCE

Clinical Psychology Intern **September 2021 – August 2022**
Patton State Hospital

- Conduct a variety of forensic evaluations (e.g., admission, malingering, competency to stand trial, psychodiagnostic, cognitive ability) consisting of a comprehensive interview, review of relevant records, and testing when appropriate
- Author reports, formulate psycholegal opinions in accordance with state statutes when necessary, and provide treatment recommendations when appropriate
- Conduct individual therapy with adults with severe mental illness utilizing evidence-based practices
- Co-facilitate Functional Rehabilitation Educational Experience (FREE) Group – a cognitive remediation, competency restoration group
- Collaborate on research proposals and projects
- Supervisors: Dr. Allison Pate, Ph.D., ABPP, Dr. Felix Sanchez, Psy.D., Dr. Cynthia Aguilar, Psy.D., Dr. David Glassmire, Ph.D., ABPP, Dr. Loren King, Ph.D., ABPP

Assistant Forensic Evaluator **September 2020 – July 2021**
Private Practice, Referrals from Texas and New Mexico

- Conducted a variety of forensic evaluations (e.g., competence to stand trial, mental state at the time of the offense, mitigation, sexual offender risk) via videoconferencing consisting primarily of a comprehensive clinical interview, review of collateral records, and testing when appropriate

- Conducted pre-employment law enforcement evaluations via videoconferencing including clinical interviewing and objective personality assessment
- Formulated psycholegal opinions with the primary supervisor in accordance with state statutes
- Co-authored reports for forensic evaluations and provided treatment recommendations when appropriate
- Supervisor: Jennifer Rockett, Ph.D.

Student Clinician and Co-Therapist **September 2020 – June 2021**
TEAM Forensic Services, Huntsville, TX

- Co-facilitated monthly, mandated, group treatment for individuals convicted of a sex offense
- Utilized manualized, evidenced-based treatment protocols
- Supervisor: Holly Miller, Ph.D., LSOTP

Assistant Forensic Evaluator **October 2018 – June 2021**
Psychological Services Center, Huntsville, TX

- Conducted court-ordered pre-trial forensic evaluations (e.g., competence to stand trial, mental state at the time of the offense, treatment recommendations) consisting primarily of a comprehensive clinical interview, review of collateral records, and testing when appropriate
- Co-authored reports for forensic evaluations, formulated psycholegal opinions with the primary supervisor in accordance with state statutes, and provided treatment recommendations when appropriate
- Supervisors: Mary Alice Conroy, Ph.D., ABPP, Wendy Elliott, Ph.D., ABPP, and Darryl Johnson, Ph.D.

Student Clinician **September 2018 – June 2021**
Psychological Services Center, Huntsville, TX

- Conducted individual therapy with adults utilizing evidence-based interventions
- Collaborated with clients on treatment planning and closely monitored treatment goals and progress
- Conducted psychological evaluations to answer referral questions (e.g., diagnostic, treatment needs)
- Engaged in suicide risk assessment and client safety planning
- Supervisors: Craig Henderson, Ph.D., Wendy Elliott, Ph.D., ABPP, Jaime Anderson, Ph.D., Temilola Salami, Ph.D., Jorge Varela, Ph.D., and Chelsea Ratcliff, Ph.D.

Student Clinician **June 2020 – September 2020**
Walker County Adult Probation, Huntsville, TX

- Conducted psychodiagnostic and substance use assessments to provide community supervision officers with treatment recommendations regarding substance use and relevant mental health needs of individuals on probation
- Co-authored reports that were utilized to determine necessary stipulations of the evaluee’s probation

- Supervisor: Darryl Johnson, Ph.D.

Student Clinician (Pre-Doctoral Therapist and Assessor) June 2019 – March 2020
Montgomery County Sherriff's Office, Conroe, TX

- Conducted brief symptomology assessments to assist jail mental health staff in determining inmates' mental health needs and relevant psychosocial history
- Engaged in long-term therapy with inmates
- Led weekly group therapy sessions focused on a variety of topics, including sleep hygiene, distress tolerance, and interpersonal effectiveness
- Provided brief interventions to inmates in acute distress
- Supervisor: Darryl Johnson, Ph.D.

DOCTORAL RESEARCH EXPERIENCE

Research Assistant August 2017 – August 2022
Forensic Psychology Research Lab
Sam Houston State University, Huntsville, TX

- Conduct statistical analyses on data for conference submissions and future publications
- Collaborate with other students in the lab on various research projects
- Perform literature reviews relevant to upcoming research projects and manuscripts
- Assist with coding and analyzing data for various projects within the lab and with researchers from other institutions (e.g. Institute of Law, Psychiatry, & Public Policy)
- Supervisor: Marcus Boccaccini, Ph.D.

Research Assistant May 2020 – September 2021
weTHRIVE Lab
Sam Houston State University, Huntsville, TX

- Assisted in a skills-based training program for law enforcement funded by the Office of the Governor
- Facilitated a modified version of Affect Regulation Training groups via a telehealth platform
- Performed various duties related to the project, including participant recruitment, lab meetings, and data collection and entry
- Participated in various projects within the lab
- Mentored masters and undergraduate students
- Supervisor: Temilola Salami, Ph.D.

Grant Funded Research Assistant October 2018 – June 2019
Institute of Law, Psychiatry & Public Policy
University of Virginia, Charlottesville, VA

- Coded over 400 Competency to Stand Trial and Mental Status at the Time of Offense evaluations
- Collaborated with faculty, staff, and students at Sam Houston and other universities

- Supervisors: Daniel Murrie, Ph.D., Angela Torres, Ph.D., Brett Gardner, Ph.D., and Marcus Boccaccini, Ph.D.

PEER REVIEWED PUBLICATIONS

Trupp, G. F., Ricardo, M. M., Boccaccini, M. T., & Murrie, D. (2021). Forensic evaluators' opinions on the use of videoconferencing technology for competency to stand trial evaluations after the onset of COVID-19. *Psychology, Public Policy, and Law.*

Ditsky, M. & **Trupp, G.** (2021). Texas Psychologists' Evaluation of Wards for Guardianship Revisited. *Texas Psychologist, 80*(1), 15-16.

Trupp, G., Preszler, J., Boccaccini, M. T., Marcus, D., Varela, J., & Turner, D. (2021). Generalizability of psychopathy network analysis findings to scores assigned for clinical practice. *Journal of Criminal Justice and Behavior*.

Reinhard, E., **Trupp, G.,** Ricardo, M. M. & Johnson, D. (2020). Competent to Work from Home? Forensic Evaluations in the Midst of a Global Pandemic. *Texas Psychologist, 79*(2), 8-10.

Boccaccini, M. T., Harris, P. B., **Trupp, G. F.,** & Varela, J. G. (2019). Diagnostic and risk assessment characteristics of offenders reevaluated for SVP civil commitment. *Journal of Sex Offender Treatment, 14*(1).

Issa, M., Falkenbach, D., **Trupp, G. F.,** Campregher, J., & Lapp, J. (2017). Psychopathy in Lebanese college students: The PPI-R considered in the context of borderline features and aggressive attitudes across sex and culture. *Journal of Personality and Individual Differences, 105*(15), 64-49.

DOCTORAL PRESENTATIONS

Trupp, G. F., Boccaccini, M. T., & Drislane, Laura (2022, March). *Clinical Utility of Psychopathy Subtypes Based on Latent Profile Analysis*. Poster presented to the Annual Meeting of the Society for the Scientific Study of Psychopathy, Virtual Conference

Trupp, G. F., Boccaccini, M. T., & Drislane, Laura (2022, March). *Clinical Utility of Psychopathy Subtypes Based on Latent Profile Analysis*. Paper presented to the Annual Meeting of the American Psychology-Law Society, Denver, CO

Yenne, E., Contreras, D., Evans, S., Partika, A., & **Trupp, G. F.** (2022, March). Computational Linguistic Analysis of Competency to Stand Trial Reports. Paper presented to the Annual Meeting of the American Psychology-Law Society, Denver, CO

Schrantz, K. N., Boccaccini, M. T., **Trupp, G. F.**, Hawes, S., & Murrie, D. C. (2020, August). *Evaluators' Use of Expressive Empathy in a Risk Assessment Interview.* Poster presented to the Annual Meeting of the American Psychological Association, Washington, DC

Trupp, G. F., Schrantz, K. N., Boccaccini, M. T., Hawes, S., & Murrie, D. C. (2020, March). *Evaluators' Attitudes, Use, and Perceptions of Empathy in a Risk Assessment Interview.* Paper presented to the Annual Meeting of the American Psychology-Law Society, New Orleans, LA

Trupp, G. F., Boccaccini, M. T., & Vera, L. (2020, March). *Triarchic Psychopathy Traits and Impression Management During a Clinical Interview.* Poster presented to the Annual Meeting of the American Psychology-Law Society, New Orleans, LA

Trupp, G. F. & Conroy, M. A. (2019, September). Competency Assessment. Presentation at Austin State Hospital.

Trupp, G. F., Preszler, J., Boccaccini, M. T., Marcus, D. J., Varela, J. G., & Turner, D. B. (2019, March). *Network Analysis of Psychopathy as Defined by the PCL-R Within the Field.* Poster presented to the Annual Meeting of the American Psychology – Law Society, Portland, OR

Schrantz, K. N., Boccaccini, M. T., Murrie, D. C., & **Trupp, G. F.** (2019, March). *Forensic evaluators' opinions regarding the use of empathy in forensic assessment.* Paper presented to the Annual Meeting of the American Psychology-Law Society, Portland, OR

Trupp, G. F. & Boccaccini, M. T. (2018, March). *Item response theory properties of PCL-R field scores*. Poster presented to the Annual Meeting of the American Psychology – Law Society, Memphis TN

DOCTORAL HONORS, AWARDS, AND CERTIFICATIONS

Texas Psychological Association Mary Alice Conroy Award Honorable Mention
November 2020

American Psychology – Law Society Student Grant in Aid ($900)
October 2020

Sam Houston State University Office of Graduate Studies Travel Fund
January 2018, January2019, January 2020

American Psychology – Law Society Outstanding Student Poster Award
March 2019

Sam Houston State University Graduate Studies Scholarship
January 2018, August 2018, January 2019, August 2019