Individual differences in the treatment of psychopathic offenders

Klein Haneveld, E.L.M.

INDIVIDUAL DIFFERENCES IN THE TREATMENT OF PSYCHOPATHIC OFFENDERS

Evelyn Klein Haneveld

Individual differences in the treatment of psychopathic offenders

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Aula der Universiteit
op woensdag 25 mei 2022, te 14.00 uur

°

door Evelyn Liesbeth Martine Klein Haneveld
geboren te Leiden

PROMOTIECOMMISSIE

Promotor:	prof. dr. J.H. Kamphuis	Universiteit van Amsterdam	III
Copromotor:	dr. W.J. Smid	De Forensische Zorgspecialisten	
Overige leden:	prof. dr. J. Hendriks	Universiteit van Amsterdam	
	prof. dr. B. Orobio de Castro	Universiteit van Amsterdam	
	dr. B.J. Verschuere	Universiteit van Amsterdam	
	prof. dr. C. de Ruiter	Universiteit Maastricht	
	prof. dr. V. de Vogel	Universiteit Maastricht	
	prof. dr. M.J.F. van der Wolf	Rijksuniversiteit Groningen	

Faculteit der Maatschappij- en Gedragswetenschappen

CONTRIBUTING AUTHORS

Jonathan D. Forbey   Professor of Psychological Science, Ball State University

Jan Henk Kamphuis   Professor of Psychological Assessment and Personality, University of Amsterdam

Dylan Molenaar   Assistant Professor of Psychometrics, University of Amsterdam

Craig S. Neumann   Professor of Clinical Psychology, University of North Texas

Wineke Smid   Head of Research Department, Van der Hoevenkliniek, Utrecht

Kelsey Timmer   Health Care Psychologist (in training), Dienst Justitiële Inrichtingen

Vivienne de Vogel   Professor of Forensic Mental Health Care, Maastricht University

Edwin Wever   Researcher, Van der Hoevenkliniek, Utrecht

LIST OF PUBLICATIONS

Chapters included in this thesis are also published as:

CHAPTER 3   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

CHAPTER 4   Klein Haneveld, E., Smid, W., Timmer, K., & Kamphuis, J. H. (2021).
            Clinical appraisals of individual differences in treatment responsivity
            among patients with psychopathy: A Consensual Qualitative
            Research study. *Criminal Justice and Behavior, 48*(8), 1031-1051.
            https://doi.org/10.1177/0093854820970597

CHAPTER 5   Klein Haneveld, E., Kamphuis, J. H., Smid, W., & Forbey, J. D.
            (2017). Using MMPI–2–RF correlates to elucidate the PCL–R
            and its four facets in a sample of male forensic psychiatric
            patients. *Journal of Personality Assessment, 99*(4), 398-407.
            https://doi.org/10.1080/00223891.2016.1228655

CHAPTER 6   Klein Haneveld, E., Molenaar, D., de Vogel, V., Smid, W., & J. H. Kamphuis
            (2021). Do we hold males and females to the same standard? A
            measurement invariance study on the Psychopathy Checklist-
            Revised. *Journal of Personality Assessment*. Advance online publication.
            https://doi.org/10.1080/00223891.2021.1947308

Voor mijn gezin
En voor mijn moeder (1933-2022)

°

"It is more important to understand the butcher than the victim."
JAVIER CERCAS

CONTENTS

x

INDIVIDUAL DIFFERENCES IN THE TREATMENT OF PSYCHOPATHIC OFFENDERS

Introduction

Psychopathy is a topic that appeals to the imagination. It is the stuff of scary films (Hannibal Lecter), disturbing documentaries (Ted Bundy), and even opera (Don Giovanni). Although Hollywood does not always get it right, in this case there is certainly a grain of truth to its sensationalistic portrayal of psychopathic individuals as dangerous predators: Psychopathy actually is a good predictor of criminal recidivism and violence. What exactly is psychopathy? The word derives from the ancient Greek words ψυχή (*psyche*, soul) and πάθος (*pathos*, suffering). In the history of classification of mental disorders, psychopathy was used as a general term for personality disorders towards the end of the 19th century (Hervé, 2007). Over the course of the 20th century, the construct was gradually narrowed down and refined to its current conceptualization. In this regard, arguably the most influential book of the 20th century is *The Mask of Sanity*, first published in 1941 by American psychiatrist Hervey Cleckley. A copy of the fifth edition (1976) can be found on the internet and is still worth reading. Based on 15 extensive and fascinating case examples, Cleckley drew up a list of 16 specific criteria for psychopathy (see Table 1). Cleckley also postulated that psychopathic individuals are not 'simply bad,' but suffer genuinely from a serious disorder. In his view, the central underlying defect is a lack or attenuation of deeply felt and integrated affective experience. In his own words:

> Without suffering or enjoying in significant degree the integrated emotional consequences of experience, the psychopath will not learn from it to modify and direct his activities as other men whom we call sane modify and direct theirs. He will lack the real driving impulses which sustain and impel others toward their various widely differing but at least subjectively important goals. He will naturally lack insight into how he differs from other men, for of course he does not differ

Table 1 *Cleckley's Psychopathic Characteristics*

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

from other men as he sees them. It is entirely impossible for him to see another person from the aspect of major affective experience, since he is blind to this order of things or blind in this mode of awareness. [...] The relatively petty states of pleasure, vexation, and animosity experienced by the psychopath have been mentioned. The opinion here maintained is that he fails to know all those more serious and deeply moving affective states which make up the tragedy and triumph of ordinary life, of life at the level of important human experience. (p. 373)

Robert Hare subsequently used Cleckley's work to develop the Psychopathy Checklist (PCL; 1980) and its successor, the Psychopathy Checklist-Revised (PCL-R; 1991, 2003). An important goal was to devise a generally accepted method of assessment that could be used to accumulate and subsequently compare research results. Psychometric analysis of the Cleckley criteria was followed by testing a list of more than 100 traits and behaviors found in the clinical literature and thought to be associated with psychopathy. Many of the items turned out to be redundant or difficult to score reliably. The PCL was constructed with the remaining items, using the 22 items which best discriminated between offenders with low and high global ratings of psychopathy and which had the strongest psychometric properties. For the subsequent revision of the PCL, in addition to

Table 2 *Items of Hare's Psychopathy Checklist-Revised (1991, 2003)*

1	Glibness/superficial charm
2	Grandiose sense of self-worth
3	Need for stimulation/proneness to boredom
4	Pathological lying
5	Conning/manipulative
6	Lack of remorse or guilt
7	Shallow affect
8	Callous/lack of empathy
9	Parasitic lifestyle
10	Poor behavioral controls
11	Promiscuous sexual behavior
12	Early behavior problems
13	Lack of realistic, long-term goals
14	Impulsivity
15	Irresponsibility
16	Failure to accept responsibility for own actions
17	Many short-term marital relationships
18	Juvenile delinquency
19	Revocation of conditional release
20	Criminal versatility

improvements to the scoring instructions, two items were dropped and the titles of quite a few items were slightly changed. See Table 2 for the items of the PCL-R that are still in use today. Hare (2003) conceptualized psychopathy as a combination of personality traits and socially deviant behaviors, and described 'the psychopath' in the following way:

> Interpersonally, psychopaths are grandiose, egocentric, manipulative, dominant, forceful, exploitative, and coldhearted. Affectively, they display shallow and labile emotions, are unable to form long-lasting bonds to people, principles, or goals, and are lacking in empathy and genuine guilt and remorse. Their lifestyle is impulsive, unstable, and sensation-seeking, they readily violate social norms and fail to fulfill social obligations and responsibilities, both explicit and implied. (p. 5)

Several self-report measures have also been developed over the past decades, such as the Levenson Self-Report Psychopathy Scale (LSRP; Levenson et al., 1995), the Psychopathic Personality Inventory (PPI/PPI-R; Lilienfeld & Widows, 2005), the Self-Report Psychopathy Scale (SRP/SRP-II/SRP-4; Hare, 1985; Paulhus et al., 2017), and the Triarchic Psychopathy Measure (TriPM; Patrick, 2010); see Sellbom and colleagues (Sellbom et al., 2018) for a recent review of these scales. Self-report measures are, of course, much easier to administer and score than a rating instrument such as the PCL-R, and interrater reliability is not a problem. However, when working with offenders in forensic psychiatry, measures based on self-report are not really a viable option. In part, this is due to conscious attempts by our patients to paint a positive picture of themselves; at other times, in my view, these patients also appear to have a genuine lack of insight into their own motives and inner world. Therefore, in this thesis we have relied on Hare's PCL-R to assess psychopathy. Although not without its weaknesses (see below), there simply is no credible alternative professional rating instrument to reliably diagnose this clinical condition. Additionally, since 2005 a PCL-R assessment is compulsory for all forensic psychiatric patients in the Netherlands with a so-called TBS order (see Chapters 3 and 4 for a description of the Dutch TBS system). Thus, the use of the PCL-R is firmly embedded in clinical practice. Because of its prominence in this thesis, and before we turn to the issue of treatment, the next paragraph is devoted to a brief overview of the instrument.

Assessment of Psychopathy with the PCL-R

The PCL-R is a rating scale that can be used by trained professionals for research and in clinical and forensic settings. The standard procedure consists of a semi-structured interview combined with extensive file and collateral information. Its 20 items can be scored 0 (*definitely does not apply*), 1 (*may apply or partly applies*), or 2 (*definitely applies*), leading to a maximum Total Score of 40. The maximum score reflects a prototypical psychopathic individual; the PCL-R thus provides a dimensional score that represents the degree to which someone matches the prototypical profile. In his 1991 manual, Hare suggested a threshold of 30 to define psychopathy; this score being approximately one standard deviation above the mean for combined North American samples of male offenders and patients. Subsequently, it appeared that in European samples a lower cut score may be more appropriate to reflect the same level of psychopathy as a score of 30 in the North American samples (Cooke & Michie, 1999; Mokros et al., 2013). In his 2003 manual

Hare already suggested that different cultures and settings (clinical, judicial or research) may call for different cut scores. Therefore, in this thesis, when applicable, we have adapted the threshold to the specific goals of the study, based on European research.

Since its construction, the PCL-R has been subjected extensively to statistical analyses to determine the underlying structure. Factor analysis initially resulted in a two-factor model of psychopathy, with Factor 1 containing the personality traits and Factor 2 the behavioral features (Hare, 1991). Subsequently, Cooke and Michie (2001) used Item Response Theory (IRT) analyses to build a new model, with three factors based on 13 of the 20 items. The remaining seven items primarily address different forms of antisocial behavior. Cooke and Michie argued that this behavior should be considered a consequence of the core features of psychopathy, as opposed to a core feature in and of itself. However, while sparking a fierce discussion about the role of antisocial behavior in psychopathy, their suggestion to adapt the PCL-R in this respect has not led to a further revision. Instead, their three factors were incorporated in identical form into the more fine-grained four-facet model of the PCL-R (Hare & Neumann, 2008; Hare et al, 2018). This model has four correlated first-order facets underpinning the higher-order two-factor model and a superordinate factor of psychopathy. See Figure 1 in Chapter 2 for a graphic depiction of this model. From a clinical perspective, the four facets point to four areas of problematic functioning. Facet 1, called 'Interpersonal', describes a glib, grandiose individual, prone to using deception to con others. Facet 2, 'Affective', contains items related to defective emotional functioning. Facet 3, 'Lifestyle', refers to a sensation-seeking, impulsive, and irresponsible lifestyle. Facet 4, 'Antisocial', contains items pertaining to aggressive and antisocial behavior from an early age onwards. In this thesis, we have mainly used the most recent four-facet model.

In addition to debate about the role of antisocial behavior, there have been several other criticisms of the PCL-R. The instrument has been accused of dominating the conceptualization of psychopathy ("the measure has become the construct") (Skeem & Cooke, 2010). This, to me, seems a little odd. In my view, there is great value in an instrument that has proven to be reliable and valid enough to be used for research as well as clinical purposes, and that has contributed significantly to our current body of knowledge about psychopathy by offering a common metric. Can the PCL-R be blamed for a possible lack of alternative instruments? More interesting is the observation that (lack of) fear and anxiety are not included explicitly in de PCL-R, whereas fearlessness and low anxiety have long been considered hallmarks of psychopathy. It is beyond the scope of this introduction to discuss the large body of research in this area, but it does not (yet) offer a clear picture (see for a review Hare, et al., 2018). Finally, it is important to understand that the PCL-R is not a dynamic instrument, in the sense that its scoring is based on lifelong patterns (as opposed to, for example, the past month or year). As a consequence, the PCL-R is not sensitive to change. Changes in behavior need to be maintained for many years, before they will affect the PCL-R score. For the evaluation of treatment effect, other outcome measures or criteria are needed or must be developed (e.g., rates of recidivism, change in risk related factors).

Treatment of Psychopathy

Whereas the literature on psychopathy is vast, and steadily growing with scores of papers, chapters, and books appearing every year, the state of the art regarding its treatment is still rather dismal. Perhaps this is best illustrated with the recently published 828-page second edition of Patrick's *Handbook of Psychopathy* (2018): Only *one* chapter, a mere 22 pages including references, is devoted to the treatment of adults *and* juveniles with psychopathy. Considering the fact that psychopathy is an important risk factor for violence and recidivism (Yang et al, 2010; Douglas et al., 2018), in combination with the fact that we have known for at least two decades that the prevalence of psychopathy in criminal justice settings is estimated to be between 15% and 25% (Hare, 1996), one would think that developing and studying treatment is a priority. However, in their review chapter in the above-mentioned Handbook of Psychopathy, Polaschek and Skeem (2018) established (as did earlier reviewers) that currently there are no randomized controlled trials, and only a handful of studies using high-quality, quasi-experimental designs, mainly with adolescents.

In fact, for many years pessimism reigned regarding the treatability of psychopathic offenders and patients. First, Cleckley (1941/1976) appears to have set the tone, concluding that no known treatment offered a cure. His idea was to focus on 'no cure but control' in specialized units, to at least protect society from chronic destructive behavior. Coming from a psychodynamic background himself, he also mentioned what appears to be a contingency management approach, suggesting a program of rewarding socially acceptable conduct and punishing irresponsible and destructive behavior. Finally, he warned of psychopathic patients simulating improvement and seducing their therapists into believing that they have truly changed. Subsequently, a retrospective study by Rice, Harris and Cormier (1992) showed that psychopathic patients actually deteriorated in treatment, which may have confirmed the belief that psychopathy is untreatable. However, several authors have commented on the specific treatment program evaluated in this study, the Oak Ridge Social Therapy Unit. Polaschek and Daly (2013) for example argued that by today's standards the way this therapeutic community operated (in the 1960's) would be considered unethical and that it may have been especially harmful for psychopathic patients. Being less compliant than other patients, they were more exposed to severe and unusual disciplinary interventions (e.g., forced administration of alcohol and drugs like LSD) and this exposure predicted recidivism across all patients. Regardless, it appears more reasonable to assume that psychopathic patients probably do not respond well to therapeutic communities of this type, than to conclude that all treatment attempts are futile.

Without a doubt, psychopathic offenders are difficult to treat. It is consistently found that psychopathic patients or offenders do not benefit as much as their non-psychopathic counterparts. When compared to non-psychopathic individuals, psychopathic patients are less motivated, show less treatment compliance, are more often involved in institutional misconduct, and have higher rates of drop-out (Ogloff et al., 1990; Wong & Hare, 2005). A relevant distinction regarding the goal of treatment, is whether the

main focus is fundamentally changing psychopathic personality or 'merely' reducing risk factors for criminal behavior. From a more ethical standpoint the former should perhaps be attempted; for the sake of society, the latter option would suffice. To my knowledge, there are no modern/recent studies addressing the effect of treatment on psychopathic personality traits, and only two that evaluate the effect of treatment of risk factors on recidivism, both from the Canadian research group of Stephen Wong. Olver and Wong (2009) reported on the therapeutic response of adult sex offenders in an intensive, high-risk program. Although psychopathy was, again, found to be a strong predictor of drop-out, those psychopaths that remained in treatment *and* made progress on risk-related treatment targets were less likely to recidivate violently. In a study with violent offenders scoring high on psychopathy, the association between improvement on risk-related treatment goals and reductions in violent recidivism was replicated (Lewis et al., 2013).

The treatment program in the two studies above was not developed for psychopathic patients specifically, but for high-risk offenders in general, incorporating the risk, need, and responsivity principles of effective correctional treatment (RNR-model; Andrews & Bonta, 2017). These principles state that forensic treatment is most effective (1) when it is matched to level of risk (higher risk implying more intensive treatment), (2) when it targets risk factors associated with reoffending (criminogenic needs), and (3) when treatment is adapted to the learning style of the offender (responsivity). It is encouraging that treatment developed for high-risk offenders in general appeared to have an effect for those psychopathic offenders that could be retained in treatment. A possible follow-up question would be whether this type of treatment can be adapted to the specific characteristics of psychopathic patients. And are all psychopathic patients the same? What, for example, distinguished the completers from the non-completers, and those who recidivated from those who did not? The notion that psychopaths form a heterogeneous group with different developmental pathways and responses to treatment actually has a long history in the clinical literature, the most well-known distinction being between primary and secondary psychopathy (Hervé, 2007). Could it be that some psychopathic profiles are more treatable than others? Is it possible that some traits make treatment especially difficult? It is these questions that inspired our research.

This Thesis

Many years ago, when I was still a psychotherapy intern, I received a training in dialectical behavior therapy (DBT). I distinctly remember the trainer explaining to us that Marsha Linehan developed DBT by sitting in an office with a one-way-mirror with a client with borderline personality disorder (BPD), and her colleagues viewing from behind the mirror. And simply trying out interventions to figure out what worked. Prior to that time, there were no evidence-based treatments for BPS; DBT was the first. I have no idea whether this was a faithful account of her *modus operandi* but it certainly appealed to my imagination at the time. Already as a student, I disliked the strife between the different schools of psychotherapy, and very much appreciated this idea of searching for mechanisms of change in clinical practice. Then, years later, I met Stephen Wong at a conference. I timidly admitted that I too would like to study the treatment of psychopathy.

His advice was: just start, it does not matter how. So I did. For this thesis, we did originally set out with a plan, but the territory was quite unknown, as it initially may have been for BPD. My supervisor Jan Henk Kamphuis allowed me to 'feel' my way through, which gave me the same sense of adventure that I imagined Linehan must have had.

The central question of this thesis is: Can individual differences be used in a systematic way to inform and improve forensic psychiatric treatment of psychopathic offenders with the goal of reducing recidivism? The thesis commences with a case presentation in Chapter 2 to lay out the land and make the case for individual differences. It was Jan Henk who came up with the luminous idea to use two famous cases: the two perpetrators from the first true crime novel *In Cold Blood*, published in 1966 by Truman Capote. Capote described the lives of the multiple homicide perpetrators in such detail that it proved possible to score the PCL-R items. When Wineke Smid and I did our consensus scoring, much to our surprise the two cases ended up with identical total scores, but different underlying facet profiles, making them quite ideal to illustrate possible individual differences among psychopathic offenders. We then used quantitative research (Latent Profile Analysis in Chapter 3) and subsequently switched to a qualitative method (Consensual Qualitative Research in Chapter 4), which I suspect is usually done the other way around. In the second study (Chapter 3) we used a large male forensic psychiatric treatment sample containing the full range of PCL-R scores to look for subtypes, and tested whether these were predictive of treatment responsivity and recidivism. In the third study (Chapter 4) we attempted to harness the wisdom of experienced clinicians to investigate which patient characteristics, treatment (provider) characteristics, and possible other factors appear to be associated with attrition. We made use of an interview and asked them to compare two male cases, both with a high level of psychopathy, but one a treatment drop-out and the other having successfully completed treatment. In the fourth study (Chapter 5) we switched to a variable-centered method (as opposed to the person-centered methods of the previous chapters), and investigated the associations between the MMPI-2-RF scale scores and the PCL-R four-facet model. Objectives were to determine how the MMPI–2–RF scales might enhance our understanding of the nature of the four facets and to discern possible implications for treatment. In the final study (Chapter 6) we originally wanted to replicate our second study in a sample of female forensic psychiatric patients, to investigate whether the same latent profiles could be found as we found with males. However, while preparing for this study we ran into doubts about the measurement invariance of the PCL-R concerning gender. Measurement invariance (MI) refers to the capacity of an instrument to measure the same underlying construct in different groups (e.g., gender, cultural background, etc.). When an instrument is not fully invariant, scores cannot be compared between the two groups. We found only one study (Bolt et al., 2004) that investigated MI in a female prison sample. While the authors concluded that the PCL-R total score could be used reliably to distinguish between psychopathic and non-psychopathic female offenders, several items were found to be non-invariant. We therefore decided to do a MI study instead, in a large sample of female forensic psychiatric patients, treated under a TBS-order. Chapter 7 provides a summary of our findings and their implications for treatment and future research.

References

Andrews, D. A., & Bonta, J. (2017). *The psychology of criminal conduct (6th ed.)*. LexisNexis/Anderson.

Bolt, D. M., Hare, R. D., Vitale, J. E., & Newman, J. P. (2004). A multigroup item response theory analysis of the Psychopathy Checklist – Revised. *Psychological Assessment*, *16*(2), 155–168. https://doi.org/10.1037/1040-3590.16.2.155

Capote, T. (1966). *In cold blood*. Penguin Books; reissued in 2009.

Cleckley, H. (1976). *The mask of sanity. An attempt to reinterpret the so-called psychopathic personality (5th ed.)*. Mosby. (Original work published 1941)

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment, 13*(2), 171-188. https://doi.org/10.1037/1040-3590.13.2.171

Douglas, K. S., Vincent, G. M., & Eden, J. F. (2018). Risk for criminal recidivism: The role of psychopathy. In C. J. Patrick (Ed.), *Handbook of psychopathy* (2nd ed., pp. 682-709). The Guilford Press.

Hare, R. D. (1980). A research scale for the assessment of psychopathy in criminal populations. *Personality and Individual Differences, 1*(2), 111–119. https://doi.org/10.1016/0191-8869(80)90028-8

Hare, R. D. (1985). Comparison of procedures for the assessment of psychopathy. *Journal of Consulting and Clinical Psychology, 53*(1), 7–16. https://doi.org/10.1037/0022-006X.53.1.7

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

Hare, R. D. (1996). Psychopathy: A clinical construct whose time has come. *Criminal Justice and Behavior, 23*(1), 25-54. https://doi.org/10.1177/0093854896023001004

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

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

Hare, R. D., Neumann, C. S., & Mokros, A. (2018). The PCL-R assessment of psychopathy: Development, properties, debates, and new directions. In C. J. Patrick (Ed.), *Handbook of psychopathy* (2nd ed., pp. 39-79). The Guilford Press.

Hervé, H. (2007). Psychopathy across the ages: A history of the Hare psychopath. In: Hervé, H., & Yuille, J. C. (Eds), *The psychopath: Theory, research, and practice* (pp. 31-56). Erlbaum.

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://psycnet.apa.org/doi/10.1037/0022-3514.68.1.151

Lewis, K., Olver, M. E., & Wong, S. C. (2013). The Violence Risk Scale: Predictive validity and linking changes in risk with violent recidivism in a sample of high-risk offenders with psychopathic traits. *Assessment, 20*(2), 150-164. https://doi.org/10.1177/1073191112441242

Lilienfeld, S. O., & Widows, M. R. (2005). *Manual for the Psychopathic Personality Inventory-Revised*. Psychological Assessment Resources.

Mokros, A., Hollerbach, P., Vohs, K., Nitschke, J., Eher, R., & Habermeyer, E. (2013). Normative data for the psychopathy checklist–revised in German-speaking countries: A meta-analysis. *Criminal Justice and Behavior, 40*(12), 1397-1412. https://doi.org/10.1177/0093854813492519

Ogloff, J. R., Wong, S., & Greenwood, A. (1990). Treating criminal psychopaths in a therapeutic community program. *Behavioral Sciences & the Law, 8*(2), 181-190. https://doi.org/10.1002/bsl.2370080210

Olver, M. E., & Wong, S. C. (2009). Therapeutic responses of psychopathic sexual offenders: Treatment attrition, therapeutic change, and long-term recidivism. *Journal of Consulting and Clinical Psychology, 77*(2), 328-336. https://doi.org/10.1037/a0015001

Patrick, C. J. (2010). *Operationalizing the triarchic conceptualization of psychopathy: Preliminary description of brief scales for assessment of boldness, meanness, and disinhibition*. Unpublished manual, Department of Psychology, Florida State University. Retrieved from http://www.phenxtoolkit.org

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

Paulhus, D. L., Neumann, C. S., & Hare, R. D. (2017). *SRP-4: Self-Report Psychopathy Scale-fourth edition*. Multi-Health Systems.

Polaschek, D. L., & Daly, T. E. (2013). Treatment and psychopathy in forensic settings. *Aggression and Violent Behavior, 18*(5), 592-603. https://doi.org/10.1016/j.avb.2013.06.003

Polaschek, D. L. L., & Skeem, J. L. (2018). Treatment of adults and juveniles with psychopathy. In C. Patrick (Ed.), *Handbook of psychopathy (2nd ed.)* (pp. 710-731). The Guilford Press.

Rice, M. E., Harris, G. T., & Cormier, C. A. (1992). An evaluation of a maximum security therapeutic community for psychopaths and other mentally disordered offenders. *Law and Human Behavior, 16*(4), 399-412. https://doi.org/10.1007/BF02352266

Sellbom, M., Lilienfeld, S. O., Fowler, K. A., & McCrary, K. L. (2018). The self-report assessment of psychopathy: Challenges, pitfalls, and promises. In C. J. Patrick (Ed.), *Handbook of psychopathy (2nd ed.)* (pp. 211–258). The Guilford Press.

Skeem, J. L., & Cooke, D. J. (2010). Is criminal behavior a central component of psychopathy? Conceptual directions for resolving the debate. *Psychological Assessment, 22*(2), 433-445. https://doi.org/10.1037/a0008512

Wong, S. C. P., & Hare, R. D. (2005). *Guidelines for a psychopathy treatment program*. Multi-Health Systems Inc.

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

Towards Personalized Medicine for Psychopathic Offenders? The Cases of Dick and Perry

Abstract

In forensic psychology, a diagnosis of psychopathy is associated with poor treatment response and a higher rate of recidivism. Evidence-based treatment aimed at reducing violent behavior does not (yet) exist. Also, there appear to be subtypes among psychopathic patients, whose treatment needs and response to therapeutic interventions may differ. Two well-known cases are presented and scored with the Psychopathy Checklist-Revised. The individuals involved have identical Total Scores but distinctive facet profiles. We will argue, that in order to improve treatment of these high-risk offenders, forensic psychology may be inspired by the concept of personalized medicine. In our view, it appears highly expedient to develop evidence-based methods to tailor treatments to key individual characteristics.

Figure 1 *Two-Factor / Four-Facet Structure of the PCL-R*

Factor 1	Factor 2
Facet 1. Interpersonal	**Facet 3. Lifestyle**
1. Glibness/Superficial Charm	3. Need for Stimulation/Proneness to Boredom
2. Grandiose Sense of Self Worth	9. Parasitic Lifestyle
4. Pathological Lying	13. Lack of Realistic, Long-Term Goals
5. Conning/Manipulative	14. Impulsivity
	15. Irresponsibility
Facet 2. Affective	**Facet 4. Antisocial**
6. Lack of Remorse of Guilt	10. Poor Behavioral Controls
7. Shallow Affect	12. Early Behavioral Problems
8. Callous/Lack of Empathy	18. Juvenile Delinquency
16. Failure to Accept Responsibility for Own Actions	19. Revocation of Conditional Release
	20. Criminal Versatility

Remaining items:
11. Promiscuous Sexual Behavior
17. Many Short-Term Marital Relationships

In the night of 15 November 1959 four members of the Clutter family were brutally murdered in their home. The two perpetrators, Dick Hickock and Perry Smith, were looking for the family safe, which was indicated to Dick by a fellow inmate. When no safe proved to be present, they killed the family and ran with less than fifty dollars in cash. Notwithstanding the cruelty of this infamous crime, it would probably by now have been forgotten, had it not been for Truman Capote's book *In Cold Blood* (1966). His description of the two killers is sufficiently elaborate to determine the level of psychopathy of both, using the Psychopathy Checklist-Revised (PCL-R; Hare, 2003). We will show that although their (high) Total Score is identical, they in fact had notably distinctive profiles that may well refer to different treatment needs and responsivity. Hence, we will use these cases to illustrate that it may be crucial to take into account individual differences in psychopathy to improve and tailor treatment components that may not necessarily cure psychopathy, but at least reduce the risk of violent recidivism.

Assessment of Psychopathy

The PCL-R (Hare, 1991, 2003) is the most widely used instrument to diagnose psychopathy in criminal justice settings. There is extensive evidence supporting its reliability and validity (Hare et al., 2018). It consists of 20 items that can be scored 0 (*definitely does not apply*), 1 (*may apply or partly applies*), or 2 (*definitely applies*). The scale can be scored by trained professionals. Scoring is based on a subject's entire life history, preferably in combination with information from a clinical interview. Most empirical research follows a cut score of 30 to classify psychopathy. The underlying structure of the PCL-R is subject to scientific debate. Exploratory factor analysis reported in the first edition of the instrument (Hare, 1991) identified two underlying factors: Factor 1 containing the personality traits typically associated with psychopathy, and Factor 2 reflecting the behavioral features. However, some argue that antisocial behavior should not be part of the diagnosis of psychopathy but should instead be regarded as a *consequence* of psychopathy (Cooke & Michie, 2001). Nevertheless, the more fine-grained Four-Facet model proposed by Hare in the second edition (2003) has been replicated repeatedly and we will adhere to this model (see Figure 1).

Treatment of Psychopathy

Psychopathy is a highly relevant clinical condition in forensic psychiatry. Its association with criminal and violent behavior is well established (Leistico et al., 2008), and its prevalence among offenders is high (Hare 1996), estimated at 15% to 25%. Specialized, evidence-based treatment does not (yet) exist; treatment programs are usually designed for high-risk offenders in general. Treatment studies typically compare the therapeutic response of psychopathic versus non-psychopathic patients. In early studies it has been shown that psychopaths are less motivated and compliant in treatment, more often involved in institutional misconduct, and have substantially higher rates of drop-out and recidivism (Ogloff et al., 1990; Wong & Hare, 2005). Findings such as these have contributed to the pessimistic view that psychopathy is very difficult, if not impossible to treat. However, other findings have been more encouraging. Olver and Wong (2009), for example, reported that, although a high level of psychopathy was a

strong predictor of drop-out, psychopaths who completed treatment and made progress on risk-related treatment targets, were less likely to recidivate violently. Hence, not all psychopaths may be equally (non-) responsive to treatment.

The notion that psychopaths form a heterogeneous group with different developmental pathways and responses to treatment actually has a long history in the clinical literature, ultimately dating back to Kraepelin (Hervé, 2007). Systematic empirical research among samples of highly psychopathic offenders is of much more recent date. We know of only two studies (Mokros et al., 2015; Olver et al., 2015), both published in 2015, that have used person-centered statistical analyses to uncover possible subtypes based on the Four-Facet structure of the PCL-R. The results of the two studies are remarkably similar. Both found two subtypes with similar profiles that do not greatly differ on the affective and lifestyle facets. Both variants are emotionally defective (Facet 2) and live aimless, irresponsible lives (Facet 3). However, one subtype, labelled 'Aggressive' by Mokros and colleagues, scored high on the antisocial facet (Facet 4) and relatively low on the interpersonal facet (Facet 1). This subtype shows persistent antisocial and aggressive behavior and appears to rely less on superficial charm and deception. The other subtype, named 'Manipulative', showed the opposite pattern: relatively high on the interpersonal facet, and lower on the antisocial facet. Let us return to the two cases from Capote's *In Cold Blood* (1966) who turned out to be a perfect illustration of these two subtypes (as depicted in Figure 2).

Figure 2 *Facet Profiles of Dick and Perry*
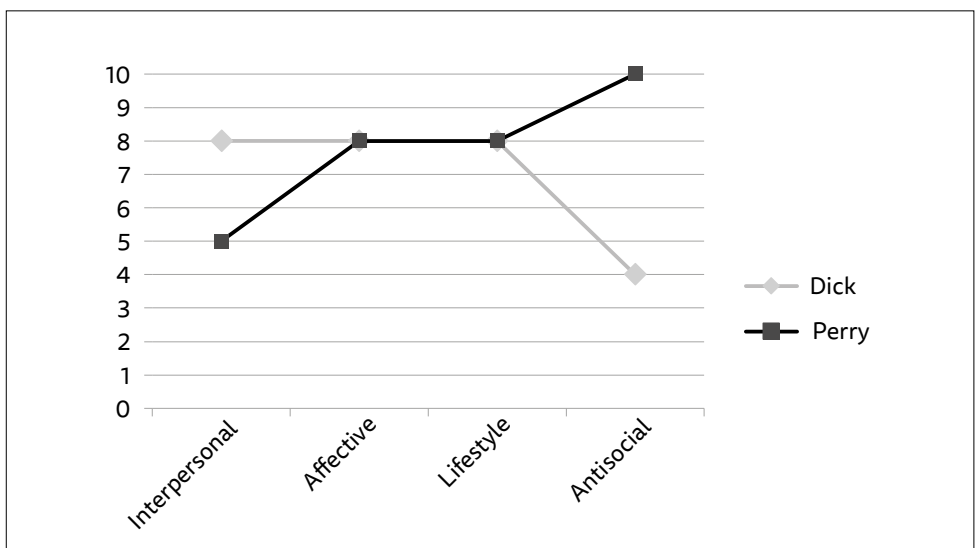
Case Descriptions

The case of Dick

Richard Eugene Hickock was born in 1931. He grew up on a modest farm in Kansas, and had one younger brother. There is almost no evidence for conduct problems prior to adolescence, except for one salient incident: he stole and destroyed all seashells belonging to a neighbor with a hammer, because he felt intensely jealous of this boy's seaside holiday. Dick was estimated to be of above average intelligence. He excelled in sports and despite motivational issues managed to graduate from high school. On the other hand, a neighbor also reported that he regularly stole things, and he was first arrested at age 17, for breaking into a drug store. After high school, Dick wanted to go to college, but his parents could not afford it. He found a job, the first of many, most of these menial jobs such as a car mechanic or painter. At age 18 he married a 15-year-old girl, against her father's will, with whom he had two children. His parents reported that from that time onward he appeared to have been continually in debt, because of reckless spending. Otherwise, his marriage appeared to be quite stable. However, Dick later reported that during these years, he first acted on his pedophiliac tendencies. Capote provided no further details, possibly because the book was first published in 1966 (a time at which this topic was taboo). But Dick also commented that when they went to the Clutter family home, "I think I went there not to rob them but to rape the girl" (Capote, 1966, p. 270). A noteworthy example of his callousness was his habit to deliberately run over dogs with his car. In 1950, Dick suffered head trauma in a car accident. According to his father, this head trauma explained his criminal tendencies, but it should be noted that his antisocial behavior significantly predated the accident. The psychiatrist who examined him in 1960 reported that he did not detect obvious signs of brain damage, but could not definitely rule it out without further examination. Around age 26 Dick started his own garage. He then got involved in an extramarital affair and neglected his business. Dick's first wife filed for divorce and he married his new lover, with whom he had his third son. He started drinking, gambling, forging cheques, then proceeded to burglary and was ultimately convicted. His second wife divorced him while he was in prison. At this time in his life, he met Perry. From another inmate, Dick had heard about the wealth of the Clutter family. He demanded to know all the details of the house (which apparently included a made up, non-existent safe) and the family. He chose to share his plan to rob them and kill any witnesses with Perry. He felt Perry was 'a natural killer'.

The Case of Perry

The early years of Perry Edward Smith, born in 1928, were of a completely different nature, characterized by brutal domestic violence and severe neglect. His parents divorced when he was seven years old. His mother was an alcoholic who had custody of both Perry and his siblings. Perry had serious behavioral problems, and was first arrested when he was eight years old. He spent the next several years in and out of child detention centers and institutions. His mother twice put him in an orphanage, where he was also maltreated because of nocturnal enuresis. Perry's father took custody of him in early adolescence. Like Dick, Perry was also of above average intelligence, and had a

passion for music, art, and poetry. However, his father did not send him to school, which Perry deeply resented and formed the basis of deep feelings of inferiority. Instead, they travelled around for several years doing odd jobs. Perry left his father after a dispute at age 16 and joined the Merchant Marine. Four years later, he joined the army. He was stationed in Japan and in Korea. He reported that he frequently had violent outbursts, and was court-martialed twice; for stealing a Japanese taxi and for demolishing a bar. After Perry completed his military service, he planned to join his father again. On his way there, he had a serious motorcycle accident, and their reunion was delayed for an entire year. The accident left Perry with substantial deformity and chronic leg pain, which led to an addiction to aspirin. When he did return to his father, their plan to start a hunting lodge failed, and his father threw him out. On his way to an 'army buddy' he was arrested for burglary, but escaped from jail. Several months later the FBI caught up with him, and he was convicted for larceny, jailbreak and car theft. It is reported that when Perry met Dick, he felt attracted to him because Dick seemed more masculine, tough, and invulnerable than he felt himself to be. To impress Dick, Perry concocted a story that he had once randomly killed a man with a bicycle chain.

Diagnostic Assessment: PCL-R Scores

The scores of both perpetrators were based on the information in Capote's book. It is important to acknowledge that this method is not entirely in accordance with the

Table 1 *PCL-R Scores of Dick and Perry*

PCL-R items	Dick	Perry
1. Glibness/superficial charm	**2**	**1**
2. Grandiose sense of self worth	2	2
3. Need for stimulation/proneness to boredom	1	1
4. Pathological lying	**2**	**0**
5. Conning/manipulative	2	2
6. Lack of remorse or guilt	2	2
7. Shallow affect	2	2
8. Callous/lack of empathy	2	2
9. Parasitic lifestyle	1	1
10. Poor Behavioral Controls	**0**	**2**
11. Promiscuous Sexual Behavior	**2**	**0**
12. Early Behavioral Problems	**0**	**2**
13. Lack of realistic, long-term goals	2	2
14. Impulsivity	2	2
15. Irresponsibility	2	2
16. Failure to accept responsibility for own actions	2	2
17. Many short-term marital relationships	**1**	**0**
18. Juvenile delinquency	**1**	**2**
19. Revocation of conditional release	2	2
20. Criminal versatility	**1**	**2**
Total Score	31	31

standard scoring procedure for the PCL-R, which is based on a clinical interview in combination with file information. Although the information compiled and presented by Truman Capote could be seen, at least partially, as file information, a clinical interview was missing. Both cases were first scored independently by two trained raters. Subsequently, a consensus score was determined (see Table 1). Figure 2 shows the PCL-R facet profiles of Dick and Perry. Both had a Total Score of 31. The facet profiles closely resemble the two subtypes found in the studies discussed above. Their Affective and Lifestyle scores (Facets 2 and 3) were identical. Both offenders suffered from severe emotional dysfunction. Throughout Capote's non-fiction novel there are many examples of shallow affect, lack of empathy and remorse. Their lives lacked direction, both before the murders as well as in the six weeks before their arrest. Decisions were often taken on the spur of the moment. Neither of them appeared to have any sense of duty or loyalty, and they had no qualms about exploiting others when possible.

Their scores on the other two facets were distinctive. With respect to Facet 1, quite a few people interviewed by Capote reported that Dick had a certain charm and was rather likeable at first sight. Even the special agents who first questioned him and were acquainted with the horrific details of the case, described him as "clean, polite, nice voice, good diction, a pretty decent-looking fellow, with a very disarming smile" (Capote, 1966, p. 210). He lied with apparent ease, and there are many instances in the book where Dick used his charm to con people. Perry did not have this ability, and was mostly a silent accomplice on their shopping sprees with bad cheques; nor was there evidence for pathological lying in his case. While his behavior and personality style did have strong appeal to certain individuals (e.g., to Capote), he did not have the same glibness as Dick. With respect to Facet 4 (the Antisocial Facet), it was Perry who scored higher. His criminal behavior was more diverse, and he was much more violent from an early age onward. The psychiatrist who had examined him described two personality features as particularly pathological: his paranoid orientation towards the world and an "ever-present, poorly controlled rage".

Finally, on the two items relating to intimate relationships and sexuality (not loading on any facet) Dick scored quite high, while Perry's did not. There is not enough information in the book to determine the nature of Dick's sexual deviancy, but it seems rather certain that it was problematic.

Discussion

Dick and Perry were sentenced to death and executed in 1965. What if they had ended up on a forensic psychiatric unit, in a high secure hospital or prison (which nowadays, at least in several European countries, would have been a realistic scenario)? Their distinctive psychopathic profiles suggest different psychological needs and potentially different responsivity, which begs the question what treatment components might have been beneficial for these different offenders. Surely, a one-size-fits-all approach would seem less than optimal. For example, aggression regulation therapy is a standard intervention in forensic treatment. In the case of our two offenders, Perry would probably need a much higher dose than Dick, having a much higher score on Facet 4 (aggressive

and antisocial behavior), but it bears mentioning that this is clinical conjecture only. In this regard, forensic treatment might seek profitable connection to personalized medicine. Recently, taking the example of oncology, Harvard psychologists Ng and Weisz (2016) reviewed what a shift to personalized intervention could mean for youth mental health. In their view it is necessary to develop evidence-based methods of assessment as well as on-going clinical decision-making to tailor treatments to key individual characteristics. The central assumption is that identification of personal characteristics predictive of individualized treatment response will ultimately improve treatment efficacy, as it did in the field of oncology. For youth mental health care, several strategies were discussed, drawing on extant knowledge about empirically supported therapies for this population.

Forensic psychiatry currently lacks the type of rich, controlled evidence base that (for example) oncology or youth mental health boast. One of several complicating factors is that (reduction of) recidivism is arguably the most important outcome variable. A 'simple' RCT, randomly assigning offenders to treatment or an alternative, and then rehabilitating them into society to see what happens, without taking into account the level of risk at discharge, is not ethically feasible. Also, small samples are often an impediment, especially when focusing on highly psychopathic patients. Accordingly, individual therapists (currently have to) follow an intuitive, personal "mix and match" strategy, with no solid evidence to back these decisions up. For example, the therapist may surmise that investing in a strong therapeutic alliance with Perry might enhance outcome, while this may prove quite treacherous with Dick. The point remains however, that these are mere conjectures lacking any empirical undergirding.

We posit that some of the strategies proposed by Ng and Weisz are now viable for transfer to forensic psychiatry. For example, existing data sets could be used to validate the two subtypes and to examine their response to treatment-as-usual. Other patient characteristics may emerge that moderate treatment effects, and could also be used to tailor treatment. Additionally, Ng and Weisz suggest the use of idiographic research designs. Single-case experiments could be employed to study the effectiveness of modular therapies developed to treat specific problem areas associated with psychopathy. In similar vein, pooling existing data sets, and use of quasi-experimental designs and propensity scoring to deal with non-random group comparisons may further the identification of profitable matching relations between individual characteristics and treatment components. We feel that examining the process of treatment with psychopathic patients in such detail may serve to elucidate what works in general, as well as what works *for whom*.

References

Capote, T. (1966). *In cold blood*. Penguin Books; reissued in 2009.

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment, 13*(2), 171-188. https://doi.org/10.1037/1040-3590.13.2.171

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

Hare, R. D. (1996). Psychopathy: A clinical construct whose time has come. *Criminal Justice and Behavior, 23*(1), 25-54. https://doi.org/10.1177/0093854896023001004

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

Hare, R. D., Neumann, C. S., & Mokros, A. (2018). The PCL-R assessment of psychopathy: Development, properties, debates, and new directions. In C. J. Patrick (Ed.), *Handbook of psychopathy* (2nd ed., pp. 39-79). The Guilford Press.

Hervé, H. (2007). Psychopathy across the ages: A history of the Hare psychopath. In: Hervé, H., & Yuille, J. C. (Eds), *The psychopath: Theory, research, and practice* (pp. 31-56). Erlbaum.

Leistico, A. M., Salekin, R. T., DeCoster, J., and 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

Mokros, A., Hare, R. D., Neumann, C. S., Pekka, S., Habermeyer, E., and Nitschke, J. (2015). Variants of psychopathy in adult male offenders: A latent profile analysis. *Journal of Abnormal Psychology, 124*(2), 372–386. https://doi.org/10.1037/abn0000042

Ng, M. Y., and Weisz, J. R. (2016). Annual research review: Building a science of personalized intervention for youth mental health. *Journal of Child Psychology and Psychiatry, 57*(3), 216-236. https://doi.org/10.1111/jcpp.12470

Ogloff, J. R., Wong, S., & Greenwood, A. (1990). Treating criminal psychopaths in a therapeutic community program. *Behavioral Sciences & the Law, 8*(2), 181-190. https://doi.org/10.1002/bsl.2370080210

Olver, M. E., Sewall, L. A., Sarty, G. E., Lewis, K., and Wong, S. C. P. (2015). 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. https://doi.org/10.1037/abn0000038

Olver, M. E., & Wong, S. C. (2009). Therapeutic responses of psychopathic sexual offenders: Treatment attrition, therapeutic change, and long-term recidivism. *Journal of Consulting and Clinical Psychology, 77*(2), 328-336. https://doi.org/10.1037/a0015001

Wong, S. C., and Hare, R. D. (2005). *Guidelines for a psychopathy treatment program*. Multi-Health Systems.

Treatment Responsiveness of Replicated Psychopathy Profiles

18

Abstract

Theory and accumulating data suggest systematic heterogeneity among offenders with psychopathic traits. Several empirical investigations converge on the nature of subtypes, but little is known about differences in treatment responsivity. We have used the four-facet model of the PCL-R to provide a framework for detecting subtypes. The present study used the full range of PCL-R scores in a sample of male violent offenders (N=190) to replicate subtypes found in a partly overlapping sample by Neumann et al. (2016), using Latent Profile Analysis (LPA), and subsequently to examine potential differences in treatment responsivity. Four subtypes emerged. Within the prototypical psychopathic group, the distinction between intent-to-treat and completers was crucial. Prototypical psychopathic offenders were significantly more likely to drop out, but completers appeared to proceed through the different phases of treatment in much the same way as the other groups. Clearly, more research is needed to elucidate treatment interfering mechanisms and their associated patient characteristics, particularly for the prototypical psychopathic group. Developing therapeutic strategies to improve treatment compliance is a necessary step in the development of specialized treatment programs for these difficult patients.

Psychopathy is a clinical syndrome characterized by a pathological personality style that is interpersonally deceptive, affectively cold, behaviorally reckless, and often overtly antisocial (Hare & Neumann, 2010). Whereas prevalence of psychopathy in society is estimated at less than 1%, in criminal justice settings, it is quite high. For example, according to Hare (1996), psychopathic offenders make up from 15% to 25% of prison populations in the US. In forensic psychiatry, psychopathy is a highly relevant syndrome because of its association with criminal and violent behavior (Leistico et al., 2008; Hare & Neumann, 2008). With regard to treatment of psychopathic offenders aimed at reducing recidivism, surprisingly little is known, considering the high estimated costs for society (e.g. Kiehl & Hoffman, 2011, estimated the cost of psychopathy at 460 billion dollars in the United States per year). An early retrospective study by Rice, Harris and Cormier (1992) suggested that treated psychopaths reoffended at a higher rate than non-treated psychopaths. Although it was noted that the therapeutic community involved in this study was an inappropriate and possibly iatrogenic treatment program (Polaschek & Daly, 2013; Reidy et al., 2013), it appears to have 'set the tone'. Arguably, pessimism dominated the field, even though several reviews concluded that there were, in fact, not enough well-designed studies to come to a conclusion about the effect of treatment on the criminal behavior of psychopaths (Salekin, 2002; D'Silva et al., 2004)

Without question, psychopathic offenders are difficult to treat. When compared to non-psychopaths, psychopaths are less motivated, show less treatment compliance, are more often involved in institutional misconduct, and have higher rates of drop-out (Ogloff et al., 1990; Wong & Hare, 2005). However, a handful of more recent studies are more encouraging. In a series of studies with adolescent offenders scoring high on the Psychopathy Checklist: Youth Version (PCL:YV; Forth et al., 2003), Caldwell and colleagues found that treatment was associated with less institutional misbehavior, and relatively slower and lower rates of violent recidivism (Caldwell et al., 2006; Caldwell et al., 2007). Olver and Wong (2009) reported on the therapeutic response of adult sex offenders in an intensive, high-risk program. Although psychopathy was (again) found to be a strong predictor of drop-out, those psychopaths that remained in treatment and made progress on risk-related treatment targets were less likely to recidivate violently. In a study from the same research group with violent offenders scoring high on psychopathy, the association between improvement on risk-related treatment goals and reductions in violent recidivism was replicated (Lewis et al., 2013).

There is almost no research comparing psychopathic offenders that do and do not recidivate. Early studies showed that a substantial portion of psychopaths is not reconvicted. For example, even with a follow-up of 8 years in the community, 20-30% of the psychopaths in a study by Serin and Amos (1995) remained free of a reconviction (for more examples, see Wong & Burt, 2007). Wong and Burt compared a group of psychopaths who recidivated with one or more violent offenses, with a group of psychopaths who did not recidivate, with a follow-up of 5 years after discharge from treatment. There was no significant difference between the two groups in mean scores on the Psychopathy Checklist-Revised (PCL-R; Hare, 1991, 2003). However, the

non-recidivating group did have significantly lower ratings on several problem areas identified by the Violence Risk Scale (VRS; Wong & Gordon, 2006). Wong and Burt conclude that psychopathic offenders differ in criminogenic needs that should be targeted in treatment, and therefore should not be seen as a homogenous group.

The acknowledgment of the heterogeneity among offenders with psychopathic traits, along with the finding that some psychopaths do seem to desist after treatment, raises the question whether treatability of psychopathy should be investigated in a more nuanced light. Two interrelated questions emerge: Is it possible to reliably identify individuals with specific psychopathic profiles? And, second, are some subtypes more treatable than others?

Theories of psychopathic variants or subtypes have quite a long history. According to Hervé (2007), Partridge first commented on possible subtypes with unique developmental pathways in a study published in 1928, but most modern authors cite Karpman (e.g. 1929, 1946), who developed a theory of psychopathic subtypes from 1929 onwards. Karpman (1946) made the distinction between primary and secondary psychopathy and described two forms of the primary subtype, i.e., an aggressive/predatory and a passive/parasitic type resulting in three subtypes (i.e., two primary and one secondary subtype). Several clinicians have since elaborated on Karpman's work (e.g. Arieti, MacCord & MacCord; see Hervé, 2007 for a review). A common thread is that primary psychopathy is thought to be more hereditary while secondary psychopathy is considered influenced by adverse early environmental factors. With respect to treatability, Karpman (1946) theorized that the primary subtypes are basically untreatable while the secondary subtype may be considerably more likely to benefit from therapy. Both hypotheses about the differences between primary and secondary psychopathy (regarding heritability and treatability) are echoed by others (e.g. Skeem et al., 2007) but seem never to have been empirically validated.

Systematic empirical research is of more recent origin. Over the past decade or so, several attempts have been made to use person-centered analytic approaches (e.g., cluster analysis, latent profile analysis) to derive subtypes among psychopathic offenders (e.g., Hicks et al., 2004; Skeem, et al., 2007). However, this research has yielded inconclusive findings as variation in mix of assessments, sample types, and psychopathy selection methods hampers accumulation of knowledge (Neumann et al., 2015). Moreover, this research offered few clues with regard to treatment response.
Several more recent investigations have employed the Four-Facet model of the Psychopathy Checklist-Revised (PCL-R; Hare, 1991, 2003) in a systematic approach to subtyping offenders. The PCL-R is probably the most widely used instrument to diagnose psychopathy, in research as well as clinical practice. The PCL-R has shown favorable psychometric properties in a large number of forensic samples (Hare et al., 2018), and has been particularly useful in assessing problematic treatment responsivity and in predicting recidivism. That said, the PCL-R is not uncontroversial in the field. Critics question whether the PCL-R accurately captures the underlying construct of psychopathy.

For example, there is an ongoing debate about whether antisocial behavior should be part of the conceptualization of psychopathy (i.e., a defining feature, as it is in the PCL-R), or whether it should be seen as (merely) a behavioral consequence of having a psychopathic personality (see for example Cooke & Michie, 2001). Nevertheless, the widespread use of the PCL-R has resulted in the accumulation of a large body of replicable findings, and the empirical support of the Four-Facet model (Hare & Neumann, 2006, 2008) offers one possible avenue towards uncovering psychopathic profiles or subtypes (Hervé, 2007; Mokros et al., 2015; Neumann et al., 2016).

Using the PCL-R, Mokros et al. (2015) conducted a latent profile analysis on a very large sample of 1451 offenders scoring 27 or higher. Three latent classes were found, which the authors named *manipulative*, *aggressive* and *sociopathic*. These appear to be similar to Karpman's two primary and secondary psychopathic subtypes. The authors suggested that the manipulative and aggressive subtypes are phenotypic variations of the 'true', primary psychopath, while the sociopathic type could be a secondary variation. The manipulative and aggressive variants both scored high on the affective and lifestyle facets. However, the manipulative group scored especially high on the interpersonal facet and relatively low on antisocial behavior, and vice versa for the aggressive variant. The authors comment that these two subtypes seem to differ primarily in the way they achieve their goals; one more through deception and manipulation, the other more through violent and intimidating behavior. The sociopathic variant scored (relatively) lower on the interpersonal facet and affective facet, but was high on the lifestyle and antisocial facets.

Another important study reported on model-based cluster analysis in 314 offenders scoring 25 or higher on the PCL-R (Olver et al., 2015). Olver et al. identified two clusters (subtypes): one scoring very high on the interpersonal and affective facets of the PCL-R, mean score 5.92 (SD = 1.06) and 6.75 (SD = 0.80) respectively, and high on the other two facets, mean score 7.17 (SD = 1.31) and 7.25 (SD = 1.40) respectively (called 'primary' by the authors); and another subtype scoring comparatively lower on the interpersonal facet, mean score 3.77 (SD = 1.17), but very high on the antisocial facet, mean score 8.57 (SD = 1.00) and high on the affective and lifestyle facets, mean score 5.67 (SD = 1.03) and 7.37 (SD = 1.19) respectively. The authors referred to this latter subtype as 'secondary' psychopathy. Note that the terms 'primary' and 'secondary', which originally referred to separate developmental pathways (hereditary versus early environmental factors), are used here in a different way, to describe different profiles of psychopathic features. For example, in the study by Mokros et al. (2015) the term 'secondary' is used to refer to a sociopathic (not truly psychopathic) variant, while Olver et al. (2015) use it to describe an aggressive variant, amply psychopathic, but lower on interpersonal features. It is outside the scope of this paper to resolve this issue, but this use of the two terms is potentially confusing.

The study by Olver and colleagues (2015) also provided the first clues about treatment response. When looking at static and dynamic risk factors of pretreatment and post treatment change on dynamic scores measured with the VRS (Wong & Gordon, 2006), Olver et al. (2015) observed that the second ('aggressive') psychopathy subtype (relatively low Interpersonal but

high on the other three facets) had higher risk and more criminogenic treatment needs at the start of treatment, and improved more during treatment. However, changes made during treatment were not related to reduced violent recidivism for this second subtype. On the other hand, the first ('manipulative' primary) subtype of psychopaths showed less improvement, but the changes they did make were indeed related to less violent recidivism. The authors speculate that the second subtype of psychopaths presented with more 'visible' emotional instability and impulsivity, and because of this profile, behavioral changes may be more easily spotted by coders of the VRS but are possibly less meaningful than changes seen in primary psychopaths. In sum, there is accumulating evidence for several distinguishable profiles among high scoring psychopathic offenders, and preliminary evidence as well as clinical conjecture suggests that these variants may be systematically related to progress in treatment and treatment outcome.

A principal limitation of these studies is that only offenders with high to very high PCL-R scores were included, effectively eliminating approximately 80% of the total sample of offenders. As several authors have noted (Mokros et al., 2015; Neumann et al., 2016), reliance on extreme groups is one particular strategy to address the question of meaningful distinctions within the offenders with psychopathic traits. Another approach is to use a comprehensive sample containing the full range of PCL-R scores. Previous research has shown that the group of offenders scoring in the medium range of the PCL-R (20-29) also has higher rates of reoffending than non-psychopathic offenders (Serin & Amos, 1995), and clinical experience suggests that this can be a very challenging group in treatment. Using the full sample and deriving a range of PCL-R facet profiles offers the opportunity to understand the critical differences, as well as similarities, between offenders with elevated PCL-R scores and those who present with mid-range scores.

Only a few studies have used person-centered subtyping methodology with samples that contained the full range of PCL-R scores. Two studies sampling male offenders both identified one non-psychopathic group and 3 clusters with differing constellations of psychopathic traits (Vassileva et al., 2005; Poythress et al., 2010). Unfortunately, both studies reported PCL-R factor (F1, F2) or facet scores only, making it impossible to determine mean total scores in the three psychopathic clusters. Nevertheless, both studies identified one 'primary' psychopathic group, one 'secondary' psychopathic group and a third group with substantial psychopathic features. In addition, Poythress et al. (2010) included several indices of treatment behavior and outcome. Approximately half of their sample (n = 660; 47%) consisted of offenders who had been court-ordered to residential drug treatment programs. Although the 'secondary' group showed less disruptive behavior during treatment than the 'primary' group, there was no significant difference in drug abuse treatment outcome or recidivism. Note, however, that the treatment involved was not based on an assessment of criminogenic treatment needs, but aimed at only one potential risk factor instead.

In more recent research, Neumann et al. (2016) conducted latent profile analyses (LPA) using four very large samples (North American and U.K. samples of male offenders, and North American and Dutch psychiatric samples; n ranging from 965 to 4865). The authors found four comparable classes across all samples: (a) a prototypical psychopathic group, (b) a group of callous-conning offenders with relatively high scores on the Interpersonal and Affective Facets, (c) a sociopathic group with relatively high scores on the Lifestyle and Antisocial Facets, and (d) a non-psychopathic general offender group. See figure 3. Noteworthy is that the proportion of, in particular, the prototypical psychopathic group (C1) varies widely between U.S. and U.K. offender samples. Possibly, varying cut scores or sampling variation (e.g. low versus high security prisons) may underlie these differences; this conjecture warrants further research. Conversely, the proportions in the forensic psychiatric samples appear to be remarkably similar. Finally, a new LPA study on male sex offenders replicated the PCL-R subtypes reported in Neumann et al. (2016), which were further validated in terms of sexual offense profiles; as expected the prototypic subtype evidenced significantly more violent sexual assaults, compared to the other three subtypes (Krstic et al., 2018).

The current study is based on a sample of 190 violent male offenders, involuntarily committed to a forensic psychiatric hospital in the Netherlands, including the full range of PCL-R scores. We used LPA to determine the number and nature of profiles of PCL-R facet scores in this population. In line with the studies by Vassileva et al. (2005), Poythress et al. (2010), and as a replication of Neumann and colleagues research (Neumann et al., 2016; Krstic et al., 2018), we expected to find one prototypical psychopathic profile, one or two profiles with substantial psychopathic traits and a non-psychopathic group. Subsequently, we related the emergent profiles to various treatment outcome variables to assess for potential differences in treatability (length of treatment phases, total treatment duration, expulsion from treatment, recidivism). In view of the limited evidence (but widely accepted clinical knowledge), we expected that patients with a prototypical psychopathic profile would have longer treatment duration, higher drop-out and higher recidivism more than other emergent profiles. We expected the non-psychopathic offender group to show most favorable outcomes. Finally, one of the potential problems with latent class analyses is that the emerging subtypes simply reflect differences in score elevation, identifying groups scoring low, moderate, and high. For this reason, we determined whether the emergent profiles predict treatment outcome over and above PCL-R Total Score.

Method

Setting and Participants
The present study was conducted at a Dutch forensic psychiatric hospital providing treatment for patients with a TBS-order. It is one of 11 facilities in the Netherlands that treat this type of patients. TBS (*ter beschikking stelling*) is a measure of mandatory intensive inpatient treatment that can be ordered by the Dutch courts, together with a sentence for violent or sexually violent offences that have a maximum sentence of at least four years. The TBS-order is imposed on those offenders whose offenses are believed

to have been associated with a psychological disorder, and who are perceived to be at risk to reoffend. The primary goal of treatment is to minimize reoffending while working towards gradual rehabilitation. Patients are admitted immediately after completing a prison sentence. Treatment cannot be terminated by either the hospital or the patient; it is only the court that has the power to extend or to terminate the TBS-order, and this evaluation takes place at least every two years. However, hospitals can decide to request for a transfer to another TBS-clinic when they feel their attempt at treatment has failed. Hence, a drop-out is not truly a drop-out but can be seen as removal from treatment.

Participants were all male patients admitted between December 2000 and November 2012 (n=192) with an indefinite TBS-order. Two were discarded because the PCL-R could not be scored (lack of sufficient information). The sample size of the current study thus consisted of the remaining 190 patients. As noted before, of this sample, 135 participants were included in the large Dutch sample (n=3224) used by Neumann et al. (2016); 55 were not included in that sample. All 190 patients were convicted of one or (frequently) several violent offenses: 146 (76.8%) patients committed a violent offence, i.e. (attempted) murder, manslaughter, or violent assault; and 68 (35.8%) patients committed a sex offence. Participants' ages ranged from 27 to 85 (M = 46.83, SD = 10.71). Ethnic constellation (based on the country of birth) was 75.8% Dutch, 11.0% Afro Caribbean, 4.7% Moroccan, 2.1% Turkish, 6.4% other.

Treatment

Although the clinic involved in this study is a high-security hospital, treatment is delivered in a generally supportive and therapeutic atmosphere. Treatment activities are varied, ranging from education and work to individual and group therapy mostly based on cognitive behavioral principles. Pharmacotherapy is included when necessary. Over the course of this study treatment methods have evolved, gradually incorporating the risk, need, and responsivity principles of effective correctional treatment (RNR-model; Andrews & Bonta, 2010). These principles state that forensic treatment is most effective (1) when it is matched to level of risk (higher risk implying more intensive treatment), (2) when it targets risk factors associated with reoffending (criminogenic needs), and (3) when treatment is adapted to the learning style of the offender (responsivity).

TBS-treatment consists of several phases with gradually increasing liberty to take leaves outside the hospital. During the first phase of treatment patients do not have permission for leave. During this phase, patients participate in an intensive daily program tailored to their personal treatment goals and individual responsivity issues. Treatment targets include risk factors like antisocial cognition, impulsivity, substance abuse, offense supportive attitudes, and lack of involvement in work and anticriminal leisure pursuits. Factors not directly related to reoffending are treated to improve responsivity (e.g. major mental disorder, social skills). Family members and friends important to the patient are contacted, screened, and when appropriate, involved in treatment. Treatment is evaluated every three months, and formal risk assessment is done at least once a year. When a sufficiently reliable working alliance has been established and a patient has made

progress on the treatment goals, additional risk assessment is done to determine whether it is feasible to start the next phase of treatment.

During the second phase patients have permission to go on supervised leave together with one or two members of the staff. Patients meanwhile continue their intensive program within the clinic. Supervised leave is mainly used to assess how the patient responds to being back in society. When the patient continues to do well, transfer to phase three is considered. During the third phase unsupervised leave is granted. For phase three the availability of a detailed relapse prevention plan is a minimum requirement. Again, structured risk assessment is done to determine whether risk of reoffending has been reduced sufficiently. During phase three various activities that were previously done within the clinic are transferred back to society (e.g. hobbies, family visits, work). Phase three basically prepares the patient for the final phase of TBS-treatment, the so-called 'transmural' phase, during which the patient lives outside the hospital but is still regularly supervised by staff members. For more information about the treatment program in the clinic involved in this study, see Binsbergen et al. (2007) or Kröger et al. (2014).

As mentioned before, a patient cannot decide himself to drop-out. It is only the staff that can make a formal request to the Dutch Ministry of Justice to transfer the patient to another hospital for another treatment attempt, or to a so called long-stay unit for a temporary suspension of treatment. In general, such a request is done for one of two reasons: (1) when after serious incidents (such as severe violence involving physical harm, or continuous deception) the staff feels the working relationship has been damaged too much, and (2) when after several years of treatment, there are insufficient gains and the staff is convinced that a patient will not benefit from further treatment at their clinic.

Permission for every form of leave (supervised, unsupervised, and transmural) also has to be acquired from the Dutch Ministry of Justice. Permission is granted only after extensive information on treatment gains has been reviewed, including the outcome of structured risk assessment. For this reason permission for leave is considered a measure of treatment success in this study.

Procedures

This research project was conducted according to the guidelines for ethical research of the Forensic Care Specialists in the Netherlands. All participants signed informed consent for the use of their file and test information.

Immediately after admission, as part of standard procedure, all patients participated in extensive psychological assessment during the first three months of treatment. The present hospital (like other forensic psychiatric institutes in the Netherlands) has separate units for personality disordered offenders and offenders with psychotic disorders and usually also adapts treatment to autistic offenders. Personality disorders were diagnosed by licensed and trained psychologists using the Dutch version of the Structured Interview for DSM-IV Personality (SIDP-IV; De Jong et al., 1996). Autism

spectrum disorders were diagnosed in consensus by a multidisciplinary team containing at least one licensed psychologist and one psychiatrist. Over the course of this study different structured screening instruments have been used to aid in diagnosing autistic disorders, alongside observations in the hospital and interviews with relatives. Psychotic disorders were diagnosed by a psychiatrist, through clinical observation and interviewing. A full DSM-IV diagnosis (including the primary diagnosis) was determined at the onset of treatment; in a few cases autistic disorders were added at a later date. In this study we have used the primary diagnoses, defined as the most prominent disorder needing treatment, to place subjects in one of four categories reflecting this disorder: 1) personality (n = 105), 2) psychotic (n = 11), 3) autistic (n = 25), and 4) personality pathology combined with psychotic episodes (n = 49). For the patients in this fourth category, the personality disorder and psychotic disorder were deemed of equal prominence, and this combination was often related to the use of drugs.

PCL-R records of all patients were independently scored by two raters, who then determined a final consensus-score together. Available file information included criminal records, police records, reports from previous institutes and from prison, and information acquired from relatives, former employers, and schools. The raters were licensed psychologists with at least a Master's Degree who had been given a three-day training in scoring the PCL-R.

Treatment phases were determined by using the official dates of permission from the Ministry of Justice as they were recorded in all patient files. In a few cases permission was temporarily withdrawn and later reinstated because of, for example, rule violation in the hospital. In these cases the first time permission was granted was used to determine the length of a treatment phase; withdrawal of permission was seen as part of the treatment process in that phase. Because the sample consists of all admissions until fairly recently (November 2012) it is important to note that only part of the sample have reached the end of treatment (n = 90). End of treatment was determined by taking the official date of discharge from the Ministry of Justice as recorded in the patient files. Treatment phases and total length of treatment were measured in months.

Recidivism was determined by inspection of the current criminal records of all patients who had a TBS-order that was terminated by the court with a follow up time of at least 12 months. New charges and convictions were counted. Charges/convictions that involved any actual, attempted, or threatened physical harm to another person were considered aggressive recidivism; convictions for offenses of a sexual nature (e.g. rape, abuse, possession of child pornography) were considered sexual recidivism. General recidivism included all charges and convictions.

Instruments

PCL-R
The *Psychopathy Checklist-Revised* (Hare, 1991, 2003) consists of 20 items, which can be scored 0 (definitely does not apply), 1 (may apply or partly applies) or 2 (definitely applies), leading to a possible maximum score of 40. Although most empirical research with

North American offenders has used a diagnostic cut score of 30 to define psychopathy, it appears that in European samples a lower cut score may be more appropriate (Mokros et al., 2013). No studies are available for Dutch samples, but based on a meta-analysis of samples from German speaking countries it was suggested that a score of 25 reflects the same level of psychopathy as a score of 30 in the North American samples (Mokros et al., 2013). Accordingly, a cut score of 25 was used in this study to define a high level of psychopathy.

 In the clinic involved in this study, the PCL-R was scored on the basis of a combination of file information and an extensive interview, as advised in the manual. Extensive psychometric properties have been documented in the manual (Hare, 2003). As noted above, raters were licensed psychologists with at least a Master's Degree who had been given a three-day training in scoring the PCL-R. The scores in this sample ranged from 3 to 38.9 (M = 23.9, SD = 7.8). The mean score on each facet was as follows: Interpersonal, 3.9 (SD = 2.2); Affective, 6.1 (SD = 1.5); Lifestyle, 6.4 (SD = 2.7); and Antisocial, 5.6 (SD = 2.8). Note that scores on the Lifestyle/Antisocial Facets are based on 5 items, while scores on the Interpersonal/Affective Facets are based on 4 items. Two items do not load on any facets but only contribute to the total score (i.e., promiscuous sexual behavior, many short-term marital relationships).
No formal inter-rater reliability estimates are available for the present sample. However, inter-rater reliability for a comparable sample from the same hospital (Hildebrand et al., 2002), largely based on the same pairs of raters, has been estimated previously. The single measure ICC was .88 for the PCL-R total score. For the classic factor 1 (comprising facet 1 and 2) the single measure ICC was .76; for factor 2 (comprising facet 3 and four of the five items of facet 4) the ICC was .83. In general, ICCs were good to excellent at the individual item-level (Mdn = .67, range .46 to .80).

SIDP-IV

The Structured Interview for DSM-IV Personality (SIDP-IV; Pfohl et al., 1997) was administered to assess DSM-IV PD symptoms. The SIDP-IV follows a topically arranged format (work, interpersonal relations, impulse control, etc.) yielding symptom scores on a 0 (absent) to 3 (strong presence) scale that are combined into the ten DSM-IV dimensional counts of PD symptoms. Its psychometric properties are well established (Widiger, 2002). Raters were extensively trained licensed psychologists.

Analytic Strategy

Latent profile analysis (LPA) is a variant of latent class analysis based on observed continuous rather than categorical variables. As a mixture-distribution model, LPA seeks to identify nominal variables that underlie the continuous data and whose identification allows de-mixing of the data (Rost, 2006). Noteworthy is that individual cases have associated probabilities for belonging to more than one latent class. The more distinct the average latent class probabilities are for the most likely class membership, the more useful will be a given latent class solution. In other words, the average probability of group membership across all subjects provides information about the quality of the

class allocation, with average probability values for viable LPA solutions generally at approximately .80 or above (Mokros et al., 2015; Rost, 2006). Both information criteria (e.g., the Bayesian information criterion-BIC) as well as modified likelihood ratio tests (LRTs; Lo et al., 2001; Nylund et al., 2007) can be used to decide on the number of latent classes. For the BIC (Schwarz, 1978), a smaller value indicates better model fit in terms of the optimal trade-off between model parsimony and residuals. Reductions in BIC value of less than 3 are considered negligible (Kass & Raftery, 1995). We primarily relied upon the BIC for gauging the best LPA solution because the bootstrap LRT is more strongly affected by nonsymmetrical data distributions (Nylund et al., 2007) and often remains inconclusive (Kupzyk, 2011). Mplus version 6.1 was used for all LPAs (Muthén, & Muthén, 2011).

As a next step, after a viable LPA solution was identified, the emergent profiles were related to various clinical outcomes. First, the profiles were related to treatment drop-out by means of Cox survival analysis, with drop-out (yes/no) as the event and recorded treatment duration as time. Treatment duration for the four PCL-R subtype profiles was compared by means of an ANOVA and, in case of a significant effect, HSD post hoc analysis. Only patients who completed treatment were included in the analysis. Next, the four PCL-R profiles were related to violent and general recidivism by means of two Cox survival analyses, with recidivism (yes/no) as the event and recorded follow up time since the end of treatment as time. Finally, all analyses were repeated while controlling for total PCL-R score, to assess the effect of subtypes over and above total score.

Results

Replication of Latent Classes
Fit indices of the LPA suggested that a four-class solution fitted the data best, see Table 1. More specifically, the BIC was lowest for four classes, and while the adjusted BIC was slightly lower for the 5-group solution, the magnitude of difference indicated that there was no improvement in model fit. For the four-group solution the average

Table 1 *Model Fit of the Latent Profile Analysis (N = 190)*

	1	2	3	4	5
Log-Likelihood	-544.76	-455.84	-436.15	-415.49	-409.22
No. of Free Parameters	8	13	18	23	28
AIC	1105.53	937.68	908.31	**876.99**	874.44
BIC	1131.51	979.90	966.76	**951.67**	965.36
Adjusted BIC	1106.17	938.72	909.74	**878.81**	876.67
Entropy	-	.84	.85	.80	.80

Note. Optimal model according to AIC, BIC, and adjusted BIC is highlighted in boldface. A five group solution produced a less than 10-point drop in AIC and adjusted BIC, which is signal that of no improvement in model fit. AIC = Akaike's information criterion; BIC = Bayesian information criterion.

probabilities for most likely latent class membership were 0.92, 0.84, 0.87, and 0.86, all of which can be considered substantial. When assigning subjects to classes based on their most likely class membership, four groups emerged: Class 1 (C1) with 77 subjects (40.5%), C2 with 32 subjects (16.8%), C3 with 56 subjects (29.5%), and C4 with 25 subjects (13.2%).

Figure 1 depicts the facet profiles of the four classes. For sake of comparability, mean item scores per facet were calculated (facets are comprised of unequal numbers of items). Mean PCL-R total scores for the four classes were: C1 31.4 (SD = 3.1, range 24.2 - 38.9), C2 17.1 (SD = 3.7, range 9.0 - 26.0), C3 22.6 (SD = 3.1, range 16.0-27.4), and C4 12.4 (SD = 4.6, range 3.0 - 19.0). With a score of 25 and higher defining a high level of psychopathy, classes C1, C2 and C3 all contain cases with moderate to strong psychopathic features. Figure 2 shows the percentages of low, medium and high total PCL-R scores within each of the C-types. The profiles shown in figure 3 show similarities and differences between the profiles reported by Neumann et al. (2016) in several much larger samples (see figure 3). As mentioned before, Neumann and colleagues chose to refer to the four subtypes as prototypic psychopaths (C1), callous-conning offenders (C2), sociopathic offenders (C3), and non-psychopathic general offenders (C4), and we followed their lead.

Figure 1 *Four LPA Subtype Profiles by PCL-R Mean Item Facet Score*

Note: C1 = prototypic psychopaths (n = 77);
C2 = callous-conning offenders (n = 32);
C3 = sociopathic offenders (n = 56);
C4 = non -psychopathic offenders (n =25).
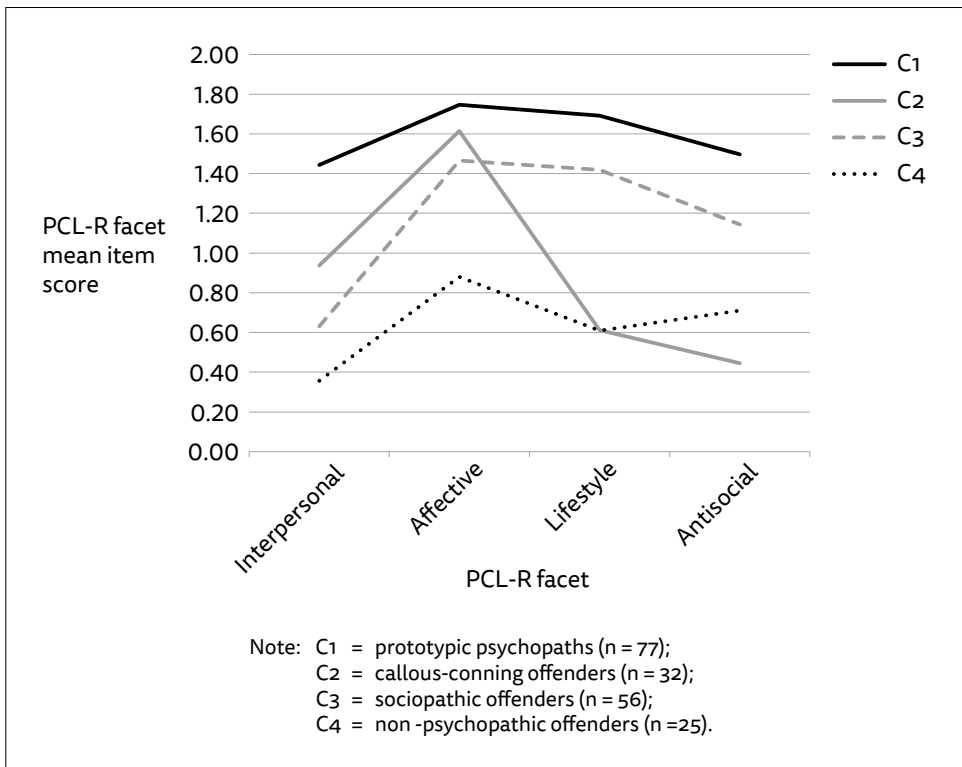
Figure 2 *Percentage of 0-15, 15-25, and >25 Total PCL-R Scores per Latent Class*

Note: C1 = prototypic psychopaths (n = 77);
       C2 = callous-conning offenders (n = 32);
       C3 = sociopathic offenders (n = 56);
       C4 = non-psychopathic offenders (n =25).
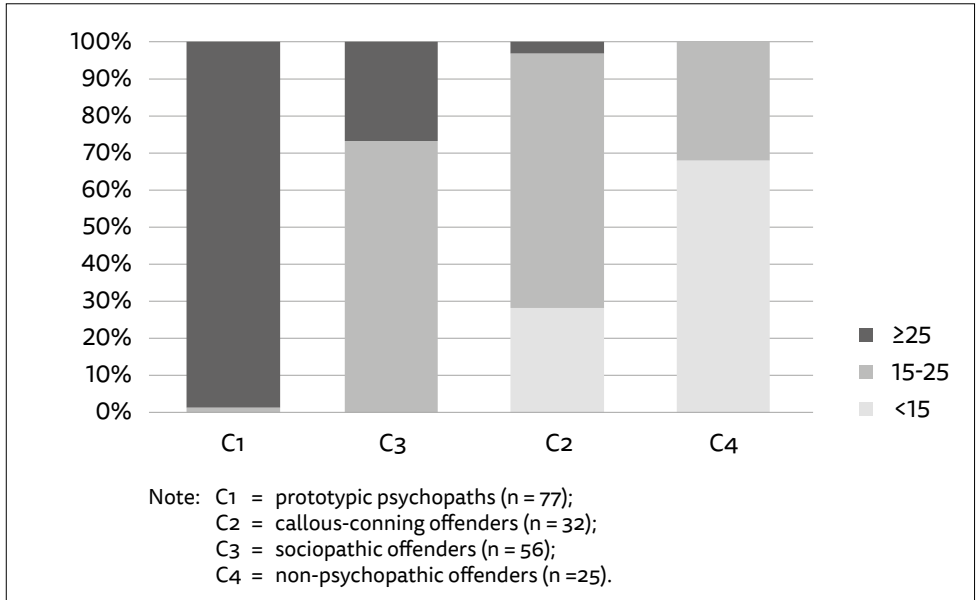
We conducted several analyses that tested aspects of subtype reliability and replicability. First, to provide information on homogeneity of the items within each facet (and thus on consistency of PCL-R item ratings within facets), we calculated mean inter-item correlations (MICs) for the four facets. The mean inter-item correlations (MICs) for the four PCL-R facets were well within the acceptable range (Clark & Watson, 1995): Interpersonal (.42), Affective (.29), Lifestyle (.48), Antisocial (.33), indicating adequate item homogeneity within each facet. We also conducted a confirmatory factor analysis (CFA) to assess fit of the Four-Facet model of psychopathy (Hare & Neumann, 2008), and obtain item factor loadings, since large factor loadings (i.e., common variance account for) are closely related to scale reliability. Consistent with a large body of findings (Neumann et al., 2015), the CFA resulted in excellent model fit (CFI = .95; RMSEA = .05), and the average of the factor loadings for all factors were strong (Interpersonal = .72; Affective = .64; Lifestyle = .77; Antisocial = 66.). Taken together, the results provide good evidence of reliability for each PCL-R facet. Second, to address subtype replicability, we compared classification accuracy across different samples. The classification accuracy for C1, C2, C3, and C4, respectively, in the current study was very much in-line with what was reported by Neumann et al. (2016), i.e., North American offenders (0.91, 0.80, 0.81, 0.86), UK offenders (0.87, 0.82, 0.87, 0.91), and North American psychiatric patients (0.84, 0.80, 0.84, 0.86). Taken together, the evidence reveals that the respective subtypes are replicated with considerable accuracy across a number of large samples.

Figure 3 *Reprinted with Permission of Neumann et al. (2016)*
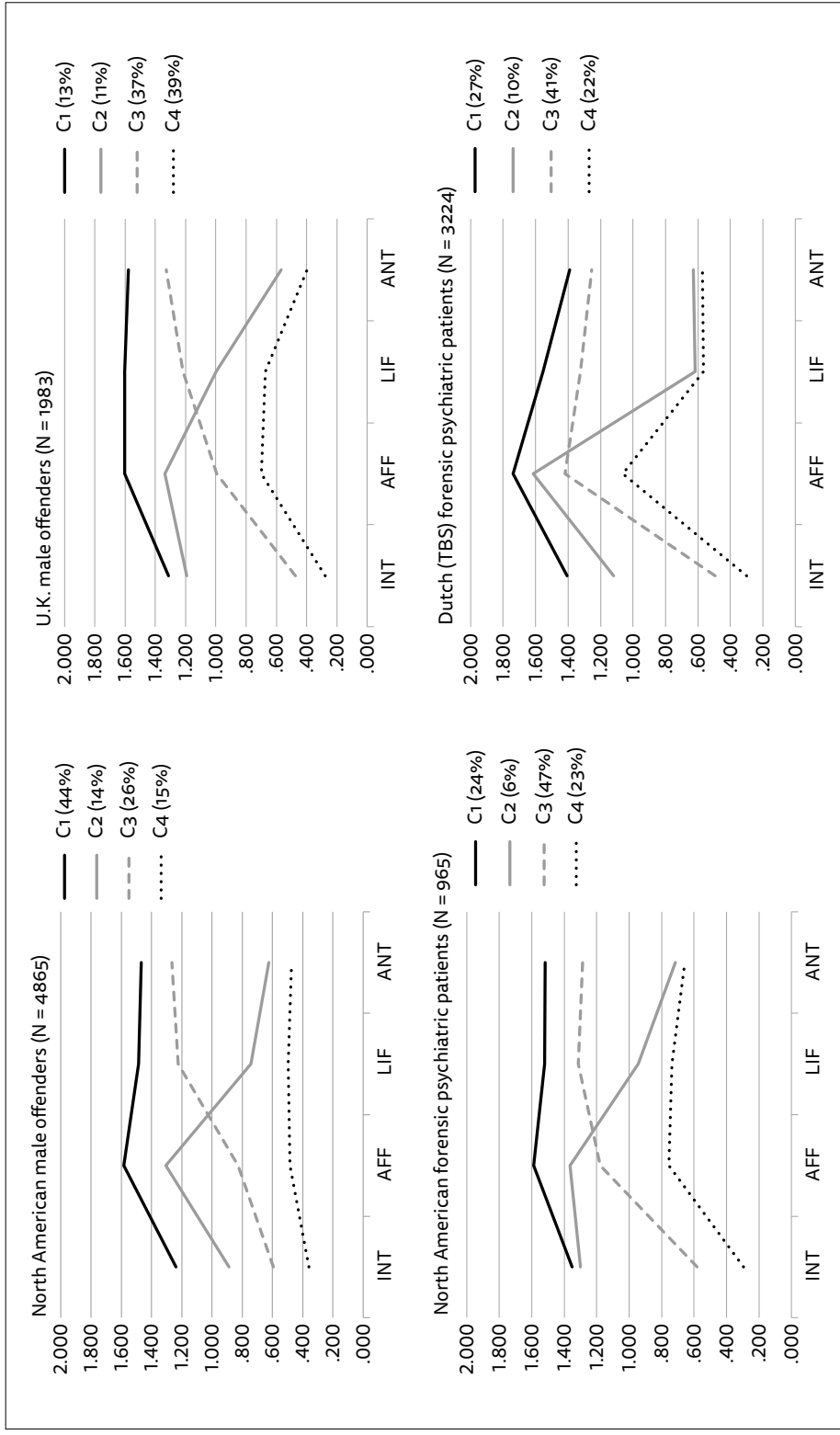
Association of Latent Classes with Treatment Variables

Table 2 contains the percentages (and number of cases) per latent class on several external variables. Nearly all subjects (97.4%) of the prototypical psychopathic group C1 were diagnosed with a personality disorder, 22.1% of whom had a comorbid psychotic disorder as well. Autism was rarely diagnosed in this group (2.6%). The other three subtypes showed considerably more variation in their primary diagnoses. Reasons for discharge also show differences between the C1 and other profiles. Treatment failed for nearly half of the offenders in C1 (46.5%), while the majority of patients in the other

Table 2 *External Variables: Percentages (numbers) per Latent Class Regarding Primary Diagnosis, Reasons for Discharge, and Recidivism*

Latent class:	C1	C2	C3	C4
PCL-R total score (M, *SD*):	31.4 (3.1)	17.1(3.7)	22.6 (3.1)	12.4 (4.6)
Variables	% (n)	% (n)	% (n)	% (n)
Primary Diagnosis				
- Personality disorder (n = 105)	75.32 (58)	56.25 (18)	33.93 (19)	40.00 (10)
- Psychotic disorder (n = 11)	0.00 (0)	12.50 (4)	7.14 (4)	12.00 (3)
- Autistic disorder (n = 25)	2.59 (2)	21.88 (7)	17.86 (10)	24.00 (6)
- Personality/psychotic (n = 49)	22.08 (17)	9.38 (3)	41.07 (23)	24.00 (6)
Total number of patients	(77)	(32)	(56)	(25)
Reasons for discharge				
- End of TBS	45.90 (28)	85.71 (18)	78.95 (30)	82.35 (14)
- Drop-out (expulsion)	46.54 (29)	9.52 (2)	18.42 (7)	11.76 (2)
- Deceased	6.56 (4)	4.76 (1)	2.63 (1)	5.88 (1)
Total number discharged	(61)	(21)	(38)	(17)
Recidivism with follow-up > 12 months				
- Aggressive	21.74 (5)	0.00 (0)	15.00 (3)	0.00 (0)
- General (including aggressive)	39.13 (9)	0.00 (0)	35.00 (7)	0.00 (0)
Total number follow-up >12 months	(23)	(14)	(20)	(13)
Length of treatment phases in months	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
- No permission for leave (n = 159)	28.09(16.33)	23.56(9.50)	23.90(13.69)	21.44(12.86)
- Supervised leave (n = 142)	17.71(9.46)	22.92(16.87)	16.986(9.46)	18.92(13.16)
- Unsupervised leave (n = 122)	19.07(16.35)	12.88(7.16)	19.37(15.36)	17.25(19.92)
- Transmural leave (n = 83)	40.55(24.84)	21.38(10.25)	29.68(16.00)	40.46(27.59)
- Total treatment duration (n = 63)	93.63[a](30.33)	70.65[b](27.80)	74.53(23.40)	88.79(27.92)

Note: C1 = prototypic psychopaths;
C2 = callous-conning offenders;
C3 = sociopathic offenders;
C4 = non-psychopathic offenders.

Note: [a] is trending longer than [b] ($p < .06$)

classes were discharged because the court terminated their TBS-order in a regular manner (C2 = 85.7%; C3 = 78.9%; C4 = 82.4%). There was no sexual recidivism in this sample. Subjects in C2 and C4 did not recidivate at all.

Clinical status of all offenders was tabulated from their records. To perform the first survival analysis, 40 subjects were marked as drop-outs. Of the 190 patients in the current sample, 53 were still in treatment at the reference date, 81 had been discharged from TBS treatment and released into society, 35 had been transferred to other TBS-clinics, 4 had been transferred to TBS long-stay facilities (civil commitment), 1 had been sent back to prison due to a new offense during treatment, 8 were deceased (of which 4 due to suicide) and 8 had been transferred to non-TBS healthcare facilities. The patients transferred to other TBS clinics, to TBS long-stay facilities (civil commitment), and to prison were considered drop-outs, the other groups were considered non-drop-outs.

Cox regression analyses indicated significantly increased explanatory power of the PCL-R profiles-model compared to baseline for drop-out (χ^2 = 21.49, df = 3, p < .000). Patients with prototypical psychopathic C1 profiles dropped out at a significantly higher rate (B = 1.66, SE = .73, Wald = 5.16, p = .02, Exp(B) = 5.27, 95% CI: 1.26 – 22.09) than the offenders with other profiles (see Figure 4). No significant differences in drop-out rates were found between the remaining profiles. Results reflected in Figure 4 suggest that the odds of surviving for C1-type offenders declined the steepest in the first two treatment phases, i.e. before unsupervised leave was granted. When PCL-R total score was entered in the Cox regression, it became the sole significant predictor of drop-out (B = .12, SE = .05, Wald = 5.17, p = .02, Exp(B) = 1.13, 95% CI: 1.02 – 1.25); C-type showed no additional significant effect.

To estimate the average duration of treatment, patients were included when they had finished treatment. Patients who left during treatment due to other reasons than finishing the treatment as intended (i.e. because of drop-out, death, transfer to a non-TBS health care facility, or being released from TBS treatment by the court against the advice of the clinic, which is very rare) were excluded. An ANOVA showed a trend regarding differences in treatment duration between the PCL-R profiles (F = 2.61, df = 3, p = .06). Post hoc tests indicated a trend (p = .09) towards longer total treatment duration for patients with a C1 profile (M = 93.63, SD = 30.33) than for those with a C2 profile (M = 70.65, SD = 27.80) (95% CI of the difference in duration: -2.68 – 48.63 months), while the other profiles fell in-between. Note that within group differences were large. The ANCOVA including total PCL-R score as a covariate, revealed no significant effect of C-type (F = 2.06, df = 3, p = .12, partial η^2 = .10).

Follow up data over a minimum period of 12 months 'at risk' was available for 70 subjects (13 - 23 for each PCL-R profile): 63 were discharged from treatment into society and 7 were transferred to non-TBS healthcare facilities. Mean follow up time was 37.96 months (SD = 22.71), ranging from 12.12 to 115.35 months. Recidivism rates were 11.4% (n = 8) for violent offenses and 22.9% (n = 16) for any/general (including violent) offenses.

Figure 4 *Survival Analyses*
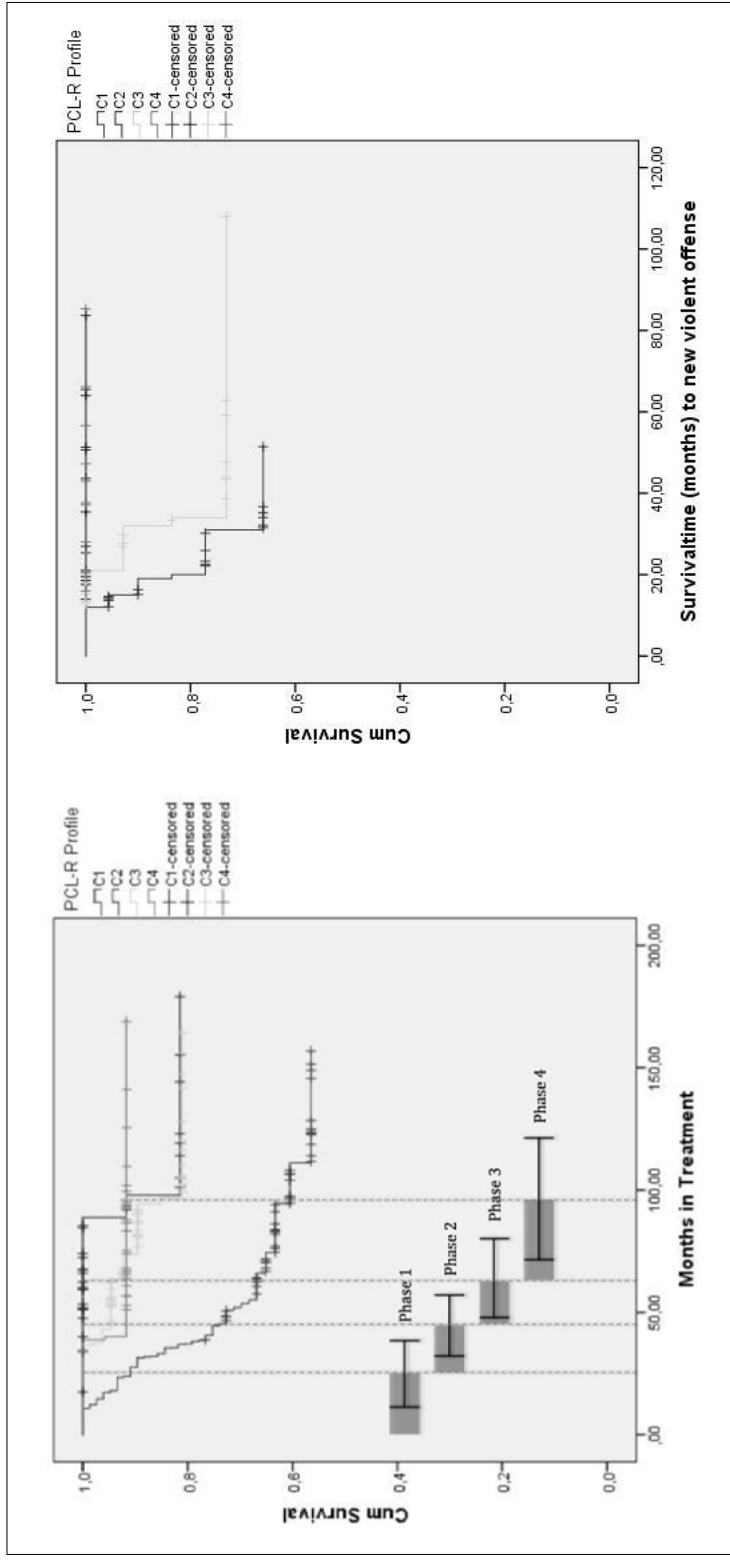
Left: Treatment drop-out failure rates among forensic psychiatric patients as a function of PCL-R profile (C1 = prototypic psychopaths (n = 77); C2 = callous-conning offenders (n = 32); C3 = sociopathic offenders (n = 56); C4 = non-psychopathic offenders (n = 25).). Including a visual indication of average treatment phase durations and standard deviations: Phase1 = no permission for leave; Phase 2 = supervised leave; Phase 3 = unsupervised leave; Phase 4 = transmural leave. *Note*: The C1 profile has a significant higher failure rate than the other profiles *Right*: Violent recidivism failure rates among forensic psychiatric patients as a function of PCL-R profile: C1 (n = 23), C2 (n = 14), C3 (n = 20), C4 (n = 13). *Note*: The PCL-R profiles-model as a whole was significantly related to violent recidivism (χ^2 = 10.65, df = 3, p = .014). Due to the low number of participants and recidivists, specifically C2 and C4, no significant differences could be assessed between the individual profiles.

Cox regression analyses indicated significantly increased explanatory power of the PCL-R profiles-model compared to baseline for both violent recidivism (χ^2 = 10.65, df = 3, p = .014) (see Figure 3) and any recidivism (χ^2 = 23.94, df = 3, p < .001). The analyses suggest that recidivism rates for patients with C1 and C3 profiles were higher than for patients with C2 and C4 profiles. However, due to the low number of participants and recidivists, the analyses could not detect significant differences between the individual profiles. A direct comparison including only the C1 and C3 profiles, showed no significant differences in terms of either violent or general recidivism.

Discussion

This study used LPA in a sample of 190 violent and/or sexually violent male offenders admitted to a forensic psychiatric hospital in the Netherlands to determine whether specific profiles could be replicated based on the four facets of the PCL-R. As expected, four distinct profiles appeared that largely replicated findings by Neumann et al. (2016) in large and diverse samples from North America (n=4,865 and n=965), the United Kingdom (n=1983), and the Netherlands (n=3224; a sample partly overlapping with the current sample). Also, the four subtypes are similar to a recent LPA study by Krstic et al. (2018), as well as those found by Vassileva et al. (2005) who used cluster analysis instead of LPA. One reviewer alerted us to the remarkably elevated facet two scores. Indeed, scores on facet two appear to be high for all three classes with psychopathic traits (see figure 1), especially when compared with the North American and U.K. offender samples (Neumann et al., 2016; see figure 3). Possibly this is due to a higher prevalence of comorbid disorders among forensic psychiatric patients, e.g. autistic spectrum disorders, intellectual disability, psychotic disorders, severe addiction. In the North American psychiatric sample (figure 3) scores on facet two also appear to be higher for the three psychopathic profiles.

Next, an attempt was made to test the clinical utility of the four emerging profiles. The first subtype (C1) scored high on all facets and had a mean total score above the cut score of 25 on the PCL-R. Their primary DSM-IV diagnosis was most often a personality disorder. This subtype appears to be *prototypically psychopathic*. The second subtype (C2) scored moderately on the Interpersonal Facet and high on Affective Facet, but much lower on Lifestyle and Antisocial Facets. Neumann and colleagues (2016) called this subtype the *callous-conning* offender, and we follow their lead. Mean total score on the PCL-R was just below 20 and this group showed a diverse range of DSM-IV classifications. The third subtype (C3) scored relatively low on the Interpersonal Facet but high on the Affective, Lifestyle and Antisocial Facets. This subtype was labeled the *sociopathic* group, and had a mean total score just above 20. It is interesting to note that the sociopathic subgroup appears to have a wider range of comorbid disorders than the more prototypical psychopathic group. One or more personality disorder(s) was the primary diagnosis in 75% of the cases, but in more than half of these cases an equally prominent psychotic disorder was diagnosed, and nearly 20% were diagnosed as autistic. The C3 group bears resemblance to the group that Olver et al. (2015) called "secondary psychopaths", presenting with higher risk and more criminogenic treatment needs at the start of treatment, compared to their "primary psychopaths". The fourth subtype, C4, scored low on all facets. This is a group of

non-psychopathic offenders who were represented in all diagnostic categories. Drop-out for C4 was low and none of the subjects in this sample recidivated.

With regard to the treatability of offenders with moderate to high psychopathic traits (classes C1, 2, and 3) the following, tentative conclusions can be drawn from this study. Within the prototypical psychopathic group, the distinction between intent-to-treat and completers seems crucial. Psychopathic offenders were significantly more likely to fail treatment, especially in the early stages of treatment before unsupervised leave was granted. In practically half of the cases clinicians prematurely transferred the offender to another hospital without finishing the program. However, psychopathic offenders who remained in treatment (completers), appeared to proceed through the different phases of treatment in much the same way as the other groups. Furthermore, prototypical psychopaths were comparable on recidivism with those assigned to the supposedly less severe, sociopathic profile (C3). In sum, although we do not know if the drop-out among the psychopathic group was in any way selective, it is possible that if they could have somehow remained in treatment, their clinical outcome could not be distinguished from the sociopathic group. At the same time, it is fair to say the C1 and C3 subtypes remain considerably dangerous given relatively high recidivism rates, compared to the C2 and C4 subtypes. Treatment durations varied between the subtypes, but also varied widely within the subgroups, with standard deviations as large as two to three years, and therefore showed no significant differences between the subgroups.

Comparing the two subtypes with moderate profiles (C2, 3) also yielded some interesting results. Whereas both profiles score mainly in the range of 15 to 25 on the PCL-R (see figure 2), callous-conning offenders (C2) appear to do much better than sociopathic offenders (C3), at least in terms of recidivism; in our sample they never recidivated. Callous and unemotional traits, although generally considered difficult to handle for treatment providers (Olver et al., 2013; Olver & Wong, 2011), probably do not warrant the concerns with regards to recidivism that they often still invoke under clinicians. This is in line with the results of a meta-analysis of the predictive validity of the PCL-R (and other instruments) with regard to violence (Yang et al., 2010). It was found that only Factor 2 (comprising of Facets 3 and 4) predicted violence.

Our final objective was to determine whether the profiles offer predictive utility for treatment, beyond simply knowing an offender's total score on the PCL-R. Unfortunately, this appears not to be the case. Classes C1, C3, and C4 most clearly appear to represent groups with different levels of profile elevations (i.e. high, moderate, low). C2 was the only possible exception, with a distinctive profile on the facets, a relatively high mean total score for European standards (17.1), and nevertheless a positive treatment outcome in terms of recidivism (no recidivism). However, when controlling for total score on the PCL-R, none of the classes showed a significant effect on any of the treatment variables in our sample.

This study has a number of limitations. First and foremost, the modest sample size and low recidivism rates ruled out several meaningful class comparisons due to low power.

The zero recidivism rates of C2 and C4 offenders interfered with the assumptions of survival analysis. And the low numbers in each of the classes may not have allowed for the potential base rate to occur. Likewise, the group of patients that had reached the transmural phase and the end of treatment via the intended route was relatively small and unequally divided over the various PCL-R profiles. Observed power was limited to detect even large effects, (.59 and .63 for transmural leave and total treatment duration, respectively). Also, this study should not be mistaken for a formal evaluation of treatment efficacy. Treatment methods evolved during the observation period of this study, and no single, formal treatment manual was followed. Rather, we focused on available indirect indices of how treatment-as-usual worked out for the emergent profiles. Thus, the current study should best be seen as a naturalistic, ecologically valid evaluation of differences in treatment responsiveness across subtypes.

Our findings provide some indications for future research. Keeping prototypical psychopathic patients in treatment appears to be a first hurdle to overcome with regard to the C1 group. As noted in the introduction, we know very little about the differences between psychopathic patients who drop out, and those that do not. And, to our knowledge, the specific treatment interfering mechanisms have never been systematically studied. One relatively accessible option to explore both differences in patient characteristics as well as treatment interfering mechanisms (apart from gathering a large dataset with exclusively high scoring offenders and more detailed information about treatment variables), would be to use a qualitative research design to generate hypotheses. These could then be tested in a series of less extensive empirical studies. Knowledge about how specific psychopathic traits affect treatability of psychopathic offenders could then inform the development of specialized treatment programs. In the RNR-model for effective forensic treatment, these traits could be seen as issues of responsivity. One example of possible traits that could influence treatability, emerged from our recent study using the Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008) to elucidate the Four-Facet model of the PCL-R (Klein Haneveld et al., 2017). Restructured Clinical scale 6 (RC6), related to interpersonal alienation, suspiciousness, and the belief that others seek to harm you, was found to be significantly correlated to the affective facet of the PCL-R. Alienation and suspiciousness are not explicitly assessed by the PCL-R. If indeed these traits are part of the problem in the treatment of psychopathic patients with high scores on Facet 2, this would call for certain adaptations of the treatment program. Klein Haneveld et al. suggested that to minimize the treatment interfering effect of distrust, treatment providers need to pay special attention to transparency about treatment methods and goals, accountability during the treatment process, and clarity about other rules and expectations.

In conclusion, although the psychopathic profiles that were replicated in this study were not found to differentiate in treatment outcome over and above PCL-R total score, we believe it remains a research priority to conduct research on traits that influence treatability of psychopathic offenders and to develop better guidelines for establishing therapeutic relationships that work with these patients.

References

Andrews, D. A., & Bonta, J. (2010). *The psychology of criminal conduct (5th ed.).* LexisNexis/Anderson.

Ben-Porath, Y. S., & Tellegen, A. (2008). *MMPI-2-RF (Minnesota Multiphasic Personality Inventory-2-Restructured Form): Manual for administration and scoring.* University of Minnesota Press.

Binsbergen, M. H. van, Keune, L. H., Gerrits, J., & Wiertsema, H. L. (2007). *Organising forensic psychiatry. Clinical practice at the Van der Hoeven Kliniek.* Forum Educatief.

Caldwell, M., Skeem, J., Salekin, R., & Van Rybroek, G. (2006). Treatment response of adolescent offenders with psychopathy features: A 2-year follow-up. *Criminal Justice and Behavior, 33*(5), 571-596. https://doi.org/10.1177/0093854806288176

Caldwell, M. F., McCormick, D. J., Umstead, D., & Van Rybroek, G. J. (2007). Evidence of treatment progress and therapeutic outcomes among adolescents with psychopathic features. *Criminal Justice and Behavior, 34*(5), 573-587. https://doi.org/10.1177/0093854806297511

Clark L. A. & Watson D. 1995 Constructing validity: basic issues in objective scale development. *Psychological Assessment 7*(3), 309–319. https://doi.org/10.1037/1040-3590.7.3.309

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment, 13*(2), 171-188. https://doi.org/10.1037/1040-3590.13.2.171

D'Silva, K., Duggan, C., & McCarthy, L. (2004). Does treatment really make psychopaths worse? A review of the evidence. *Journal of Personality Disorders, 18*(2), 163-177. https://doi.org/10.1521/pedi.18.2.163.32775

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

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

Hare, R. D. (1996). Psychopathy: A clinical construct whose time has come. *Criminal Justice and Behavior, 23*(1), 25-54. https://doi.org/10.1177/0093854896023001004

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

Hare, R. D., & Neumann, C. S. (2006). The PCL-R assessment of psychopathy. In: C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 58-88). The Guilford Press.

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

Hare, R. D., & Neumann, C. S. (2010). The role of antisociality in the psychopathy construct: comment on Skeem and Cooke. *Psychological Assessment, 22*(2), 446-454. https://doi.org/10.1037/a0013635

Hare, R. D., Neumann, C. S., & Mokros, A. (2018). The PCL-R assessment of psychopathy: Development, properties, debates, and new directions. In C. J. Patrick (Ed.), *Handbook of psychopathy* (2nd ed., pp. 39-79). The Guilford Press.

Hervé, H. (2007). Psychopathy across the ages: A history of the Hare psychopath. In: Hervé, H., & Yuille, J. C. (Eds), *The psychopath: Theory, research, and practice* (pp. 31-56). Erlbaum.

Hicks, B. M., Markon, K. E., Patrick, C. J., Krueger, R. F., & Newman, J. P. (2004). Identifying psychopathy subtypes on the basis of personality structure. *Psychological Assessment, 16*(3), 276-288. https://doi.org/10.1037/1040-3590.16.3.276

Hildebrand, M., de Ruiter, C., de Vogel, V., & van der Wolf, P. (2002). Reliability and factor structure of the Dutch language version of Hare's Psychopathy Checklist-Revised. *International Journal of Forensic Mental Health, 1*(2), 139-154. https://doi.org/10.1080/14999013.2002.10471169

Jong, C.A.J. de, Derks, F.C.H., Oel, C.J. van & Rinne, Th. (1996). *SIDP-IV: Gestructureerd interview voor de DSM-IV persoonlijkheidsstoornissen* [*SIDP-IV: Structured Interview for DSM-IV Personality*]. Stichting Verslavingszorg Oost Brabant.

Karpman, B. (1929). The problem of psychopathies. *Psychiatric Quarterly, 3*, 495-525. https://doi.org/10.1007/BF01585441

Karpman, B. (1946). Psychopathy in the scheme of human typology. *The Journal of Nervous and Mental Disease, 103*(3), 276-288. https://psycnet.apa.org/doi/10.1097/00005053-194603000-00007

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*(430), 773-795. https://doi.org/10.2307/2291091

Kiehl, K. A., & Hoffman, M. B. (2011). The criminal psychopath: History, neuroscience, treatment, and economics. *Jurimetrics, 51*, 355.

Klein Haneveld, E., Kamphuis, J. H., Smid, W., & Forbey, J. D. (2017). Using MMPI–2–RF correlates to elucidate the PCL–R and its four facets in a sample of male forensic psychiatric patients. *Journal of Personality Assessment, 99*(4), 398-407. https://doi.org/10.1080/00223891.2016.1228655

Kröger, U., Beek, D. van, Wolf, P van der, Klein Haneveld, E., Geest, H. van & Geraerts, R. (2014). *Treatment of psychopathy: A mission impossible?* Forum Educatief.

Krstic, S., Neumann, C. S., Roy, S., Robertson, C., 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

Kupzyk, K. A. (2011). *Introduction to mixture modeling.* Retrieved from https://cyfs.unl.edu/cyfsprojects/videoPPT/84329ab252efffb78f4daf28ecbf6274/110121-Kupzyk.pdf

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*, 28-45. https://doi.org/10.1007/s10979-007-9096-6

Lewis, K., Olver, M. E., & Wong, S. C. (2013). The Violence Risk Scale: Predictive validity and linking changes in risk with violent recidivism in a sample of high-risk offenders with psychopathic traits. *Assessment, 20*(2), 150-164. https://doi.org/10.1177/1073191112441242

Lo, Y., Mendell, N., & Rubin, D. (2001). Testing the number of components in a normal mixture. *Biometrika, 88*(3), 767–778. https://doi.org/10.1093/biomet/88.3.767

Mokros, A., Hare, R. D., Neumann, C. S., Santtila, P., Habermeyer, E., & Nitschke, J. (2015). Variants of psychopathy in adult male offenders: A latent profile analysis. *Journal of Abnormal Psychology, 124*(2), 372-386. https://doi.org/10.1037/abn0000042

Mokros, A., Hollerbach, P., Vohs, K., Nitschke, J., Eher, R., & Habermeyer, E. (2013). Normative data for the Psychopathy Checklist–Revised in German-speaking countries: A meta-analysis. *Criminal Justice and Behavior, 40*(12), 1397-1412. https://doi.org/10.1177/0093854813492519

Muthén, L. K., & Muthén, B. O. (2011). *Mplus (Version 6.12 for Mac)* [Computer software]. Author.

Neumann, C. S., Hare, R. D., & Pardini, D. A. (2015). Antisociality and the construct of psychopathy: Data from across the globe. *Journal of Personality, 83*(6), 678-692. https://doi.org/10.1111/jopy.12127

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* (pp.14-31). Routledge.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14*(4), 535–569. https://doi.org/10.1080/10705510701575396

Ogloff, J. R., Wong, S., & Greenwood, A. (1990). Treating criminal psychopaths in a therapeutic community program. *Behavioral Sciences & the Law, 8*(2), 181-190. https://doi.org/10.1002/bsl.2370080210

Olver, M. E., Lewis, K., & Wong, S. C. (2013). Risk reduction treatment of high-risk psychopathic offenders: The relationship of psychopathy and treatment change to violent recidivism. *Personality Disorders: Theory, Research, and Treatment, 4*(2), 160-167. https://doi.org/10.1037/a0029769

Olver, M. E., & Wong, S. C. (2009). Therapeutic responses of psychopathic sexual offenders: Treatment attrition, therapeutic change, and long-term recidivism. *Journal of Consulting and Clinical Psychology, 77*(2), 328-336. https://doi.org/10.1037/a0015001

Olver, M. E., & Wong, S. (2011). Predictors of sex offender treatment dropout: Psychopathy, sex offender risk, and responsivity implications. *Psychology, Crime & Law, 17*(5), 457-471. https://doi.org/10.1080/10683160903318876

Olver, M. E., Sewall, L. A., Sarty, G. E., Lewis, K., and Wong, S. C. P. (2015). 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. https://doi.org/10.1037/abn0000038

Pfohl, B., Blum, N., & Zimmerman, M. (1997). *Structured Interview for DSM-IV Personality (SIDP-IV)*. American Psychiatric Press.

Polaschek, D. L., & Daly, T. E. (2013). Treatment and psychopathy in forensic settings. *Aggression and Violent Behavior, 18*(5), 592-603. https://doi.org/10.1016/j.avb.2013.06.003

Poythress, N. G., Edens, J. F., Skeem, J. L., Lilienfeld, S. O., Douglas, K. S., Frick, P. J., Patrick, C. J., Epstein, M., & Wang, T. (2010). Identifying subtypes among offenders with antisocial personality disorder: A cluster-analytic study. *Journal of Abnormal Psychology, 119*(2), 389-400. https://doi.org/10.1037/a0018611

Reidy, D. E., Kearns, M. C., & DeGue, S. (2013). Reducing psychopathic violence: A review of the treatment literature. *Aggression and Violent Behavior, 18*(5), 527-538. https://doi.org/10.1016/j.avb.2013.07.008

Rice, M. E., Harris, G. T., & Cormier, C. A. (1992). An evaluation of a maximum security therapeutic community for psychopaths and other mentally disordered offenders. *Law and human behavior, 16*(4), 399-412. https://doi.org/10.1007/BF02352266

Rost, J. (2006). Latent-Class-Analyse [Latent class analysis]. In F. Petermann & M. Eid (Eds.), *Handbuch der psychologischen diagnostik* [*Handbook of psychological assessment*]. Hogrefe.

Salekin, R. T. (2002). Psychopathy and therapeutic pessimism: Clinical lore or clinical reality? *Clinical Psychology Review, 22*(1), 79-112. https://doi.org/10.1016/S0272-7358(01)00083-6

Serin, R. C., & Amos, N. L. (1995). The role of psychopathy in the assessment of dangerousness. *International Journal of Law and Psychiatry, 18*(2), 231-238. https://doi.org/10.1016/0160-2527(95)00008-6

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

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

Vassileva, J., Kosson, D. S., Abramowitz, C., & Conrod, P. (2005). Psychopathy versus psychopathies in classifying criminal offenders. *Legal and Criminological Psychology, 10*(1), 27-43. https://doi.org/10.1348/135532504X15376

Widiger, T. A. (2002). Personality disorders. In: M. M. Antony & D. H. Barlow (Eds.), *Handbook of assessment and treatment planning for psychological disorders.* The Guilford Press

Wong, S. C., & Burt, G. (2007). The heterogeneity of incarcerated psychopaths: Differences in risk, need, recidivism, and management approaches. Hervé, H., & Yuille, J. C. (Eds.), *The psychopath: Theory, research, and practice.* Erlbaum.

Wong, S. C., & Gordon, A. (2006). The validity and reliability of the Violence Risk Scale: A treatment-friendly violence risk assessment tool. *Psychology, Public Policy, and Law, 12*(3), 279-309. https://doi.org/10.1037/1076-8971.12.3.279

Wong, S. C. P., & Hare, R. D. (2005). *Guidelines for a Psychopathy Treatment Program.* Multi-Health Systems.

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

Clinical Appraisals of Individual Differences in Treatment Responsivity Among Psychopathic Patients: A Consensual Qualitative Research Study

Abstract

This study addressed which factors expert clinicians consider crucial in successful completion versus drop-out in the mandatory forensic psychiatric treatment of psychopathic patients in the Netherlands. 11 clinicians were interviewed, about patient characteristics, treatment (provider) characteristics, and other factors they deemed associated with failure (transfer to another facility) or completion. The interviews were coded using the guidelines of Consensual Qualitative Research (CQR; Hill, 2012). Overall, extremely high scores on PCL-R Facets 1 (deceitful interpersonal style) and 2 (defective affective experience) were thought to impede treatment retention, particularly by their negative impact on motivation and therapeutic relationship. Older patients, those with a prosocial network, and/or patients with comorbid borderline traits appeared to fare better. Treatment success was deemed more likely when treatment goals and expectations are stipulated in a concrete fashion, when an extended and gradual resocialization trajectory is offered, and the treatment team is expert, cohesive, and stable.

Both clinical experience and empirical research into inpatient forensic psychiatric treatment for psychopathic offenders to date suggest that these patients are particularly difficult to treat. In early treatment studies, psychopathic offenders showed less motivation, less compliance with treatment, and conversely, more often involvement in institutional misconduct, and more drop-out as compared to non-psychopathic offenders (Ogloff et al., 1990; Wong & Hare, 2005). However, a handful of more recent reviews is more encouraging (e.g. Hecht et al., 2018; Polaschek & Skeem, 2018; Polaschek, 2014; Reidy et al., 2013). Moreover, these early findings do not rule out that for some subgroups of psychopathic offenders, treatment may be effective in reducing criminal behavior. For example, Olver and Wong (2009) reported on the therapeutic response of sex offenders in a high-risk program. Psychopathy was found to be a strong predictor of drop-out. However, those psychopaths who remained in treatment and made progress on risk-related treatment targets were less likely to recidivate violently than the non-completers. In similar vein, a Dutch study (Klein Haneveld et al., 2018) targeted individual differences in treatability by seeking empirically based subtypes of male violent offenders in a forensic psychiatric hospital. Based on latent profile analysis of scores on the Psychopathy Checklist-Revised (PCL-R; Hare, 2003), three psychopathic profiles emerged: one prototypical group, with high PCL-R scores (and thus displaying virtually all defining features of this composite construct), and two groups with moderate psychopathic scores on the PCL-R. Drop-out was (again) very high in the prototypical psychopathic group (nearly 50%), but those who remained in treatment showed recidivism rates commensurate with the less severe psychopathic group. Hence, one hurdle to overcome in the search for effective treatments for psychopathic offenders is the issue of attrition.

To our knowledge, only a handful of studies that have examined treatment attrition in psychopaths. These studies (Olver & Wong, 2011; Cullen et al., 2011; Jeandarme et al., 2017; Sewall & Olver, 2019) reported on the association between non-completion and the Two-Factor or Four-Facet models of the PCL-R (Hare, 1991, 2003), or the Two-Factor model of the screening version (PCL:SV; Hart et al., 1995). In a group of 154 sex offenders receiving a high intensity treatment program, Olver and Wong (2011) found that all facets were related to non-completion, but that only the Affective Facet (Facet 2) made a significant unique contribution to the prediction of drop-out. Olver and Wong suggested that shallow affect, callousness, and lack of empathy interfere with the ability to 'connect' with the patient and to form a working alliance in treatment. In a British study (Cullen et al., 2011) in 84 mentally disordered offenders sampled from six medium security forensic hospitals, drop-out was studied as part of a randomized controlled trial of a cognitive skills program. In this group, Factor 2 (impulsive and antisocial behavior) was found to predict treatment attrition. A more recent Belgian study also assessed outcome in mentally disordered offenders receiving medium security treatment (N = 224). Contrary to the British study, Factor 1 predicted drop-out after controlling for offense-related and clinical variables (Jeandarme et al., 2017). Factor 1 encompasses Facet 1 (Interpersonal) and 2 (Affective). Finally, in a study by Sewall & Olver (2019) in 302 sex offenders, only Facet 3 (Lifestyle) made a significant unique contribution to the prediction of drop-out,

and a trend was noted for the Affective Facet 2 (*p* = .081). In sum, although the Affective Facet emerges as the only replicated predictor, no clear picture emerges from these studies about which aspects of psychopathy, as measured with the PCL-instruments, are associated with treatment attrition among offenders. Moreover, all studies focused exclusively on offender characteristics associated with drop-out. Clinical experience (Kröger et al., 2014) strongly suggests that contextual factors may also play a role, such as characteristics of the treatment team, or institutional policies unrelated to either the offender or the treatment program. We are not aware of any studies reporting on these possible determinants of treatment drop-out versus completion in psychopathic offenders.

The Present Study

Our goal was to harness clinical experience and to generate more inclusive hypotheses on differences in treatability among psychopathic offenders. In view of the current state of empirical research, we decided to use a qualitative research design. This approach made it relatively straightforward to explore a wider range of possible factors related to drop-out beyond offender characteristics only. In our analytic approach, we adhered to the guidelines provided by the Consensual Qualitative Research paradigm (CQR; Hill et al., 1997; Hill et al., 2005; Hill, 2012). CQR is characterized by (a) the use of semi-structured interviews; (b) a method of data analysis in which a primary team of several judges bring multiple perspectives to the data and then work towards consensus about the meaning of the data; (c) a coding system that consists of domains, core ideas, and categories that are used in the cross-analysis; and (d) at least one auditor who checks the work of the primary team. We focused on the following questions:

1) Which patient characteristics are associated with failure or successful completion of treatment?
2) Which factors related to the treatment providers are associated with failure or successful completion of treatment?
3) Which other factors influence failure or successful completion of treatment?

Method

Setting

In the Netherlands, offenders who were deemed psychopathic have been committed for forensic psychiatric treatment since 1928. In that year the so-called "Psychopath Laws" came into existence, which made it possible to impose a TBS-order for mentally disordered offenders. TBS (*ter beschikking stelling*) is a measure of mandatory intensive treatment, usually of indefinite length, that can be ordered by the Dutch courts, together with a sentence for violent or sexual offenses. The primary goal of treatment is to minimize risk of reoffending while working towards gradual rehabilitation. Notably, assessment and treatment methods have varied widely over the decades. The translation of the Psychopathy Checklist-Revised into Dutch in 2001 (PCL-R; Vertommen et al., 2002) marked the beginning of systematic assessment of psychopathy. Since 2005, the Dutch Ministry of Justice requires a PCL-R assessment for every TBS-patient in the Netherlands

(if sufficient information is available). The result of this regulation is that the concept of psychopathy as defined by the PCL-R has become embedded in clinical practice.

Participants

The present study was conducted at two of eleven high-security hospitals in the Netherlands providing treatment for patients with a TBS-order. Hill et al. (1997) suggested that expert interviewees be selected from a homogeneous population that is very knowledgeable about the subject under investigation. For this reason, all participants were licensed psychologists with at least a Master's Degree and a minimum of five years of post-license clinical experience in a specialized TBS hospital. We were confident that these requirements would result in a sufficiently knowledgeable and homogeneous sample. The first ten interviewees (six of whom were female) were employed in the same clinic. During the analysis of the data and the development of the coding system, a licensed psychologist (male) from a second clinic was asked to participate, to determine whether the coding system was applicable to his interview as well, and to ascertain whether new themes emerged (which was not the case).

Treatment

All patients, for whom treatment progress was evaluated in the interviews, were admitted to a high secure forensic psychiatric hospital under a TBS-order. Offenders are admitted immediately after completing a prison sentence. TBS treatment starts with a high security inpatient phase. When an independent board of professionals assesses that sufficient progress has been made on risk-related targets of treatment, permission for leave outside the hospital can be obtained from the Ministry of Justice; first supervised, then unsupervised leave. In the final phase, patients leave the hospital but remain under supervision until their eventual release. All decisions about leave are based on an extensive review of the treatment gains, including structured risk assessment. TBS treatment currently takes seven to eight years on average, including several years outside the TBS clinic (Tbs inzichtelijk, 2019). The court evaluates progress in treatment every two years, and decides whether or not to extend or terminate the order. In this study, treatment was considered successful when a patient proceeded through all the phases of treatment, and the court subsequently ended the TBS-order. Because it is impossible to drop out of the (mandatory) TBS treatment, a patient was considered a drop-out (or treatment failure) when the forensic psychiatric treatment failed or stalled and a transfer to another hospital for alternative treatment was deemed necessary.

Research Team

The research team consisted of four researchers with varying backgrounds and prior experience: one senior researcher specialized in forensic psychiatry, a psychology student working on her Master's Degree, a licensed clinical psychologist/psychotherapist with extensive experience in working with psychopathic offenders (first author), and a full-time university professor/ clinical psychologist with substantial forensic research experience. Three were female, and all were native Dutch. The first three formed the primary team, the university professor served as the auditor. As advised by Sim et al. (2012), the team

spent considerable time discussing their expectations and potential biases during the first meeting. On occasion, these biases would surface during the primary research team data analysis efforts (e.g., a mildly cynical reaction to an example of patient behavior while formulating a core idea). When this happened, these biases were immediately discussed to limit their influence. Nevertheless, it is important to acknowledge that the members of the research team were co-constructors of the data and the results. A different team may have come to slightly different results; see for example a comparison of cross-analyses by two teams with the same data by Ladany and colleagues (Ladany et al., 2012).

Procedures

Recruiting Participants
This research project was approved by the University of Amsterdam faculty of Social and Behavioral Sciences ethics review board. Expert psychologists were recruited per email by the first author. The first author knew all participating psychologists as peer colleagues. Of the fourteen experienced psychologists who were invited for an interview, ten consented to participation. All participants gave written informed consent for the recording of the interview and the use of the transcript.
Ideally, in qualitative research, data collection continues until no new themes emerge. In an update of CQR, Hill and her colleagues (Hill et al., 2005) argued that it has proven difficult to establish with certainty whether this has been achieved. Based on 25 CQR based studies, they recommend a sample of 8-15 participants, depending on the homogeneity of the group (i.e., with a heterogeneous sample requiring more participants). Given our selection criteria (see under "Participants"), we were confident that our sample of eleven experienced forensic practitioners was sufficiently homogeneous and expert to produce consistent results.

Interviewing
The first author conducted the interviews between November 2017 and April 2018. All interviews were conducted by telephone and were recorded using the 'TapeACall' app. In two cases, part of the respective interview did not record well. Because the interviewer also took extensive interview notes, the missing information could be recovered from these notes (the same day).

In line with recommendations by Burkard et al. (2012), the instructions and the interview questions were sent to the participants prior to the interview (see Table 1). Accordingly, participants can more fully appraise to what they may give their informed consent (or not). Considering the fact that all participants knew the interviewer as a colleague, this also enabled them to consider whether they felt free to discuss their thoughts on this particular subject with her specifically. Additionally, having the protocol in advance gave participants time to reflect on their complex experiences with these difficult patients. We asked participants to select two male patients whom they had treated in the past ten years, both scoring 28 or higher on the PCL-R, with one of the patients having completed treatment as intended (i.e., completion), and the other having been transferred to another

Table 1 *Semi-Structured Interview*

Instructions for the preparation of the interview:

- Please select two male patients whom you have treated in the past 10 years, both scoring 28 or higher on the PCL-R. It is important that one of the patients completed treatment as intended, while the other was transferred to another clinic due to a failed treatment attempt.
- See the interview questions below. For both cases, the same questions will be asked. To prepare for the interview, please take one of the cases in mind, read the questions, and reflect upon your answers. Then repeat the same procedure for the second case.

Interview schedule:

1. Which specific patient characteristics of the case you have chosen are associated with the success/failure of treatment?
2. For each characteristic, can you describe a specific situation or event that illustrates how this characteristic has influenced the course of treatment?
3. Which factors related to the treatment team are associated with the course of treatment?
4. Can you think of a specific example that illustrates how each factor influences treatment?
5. According to you, are there other factors not yet discussed on the basis of these two cases, which may be essential to take into account when treating psychopathic offenders? Can you illustrate these factors with specific examples?

clinic due to a failed treatment attempt (i.e., drop-out). We felt that discussing two specific cases might serve to elicit more specific and concrete answers, rather than vague generalities.

Finally, several steps were taken to guarantee the anonymity of the cases that were discussed by the participants during the interview. Participants were only asked for the PCL-R total score and the year of birth of the patients involved. When accidentally a name was mentioned, we deleted these from the transcripts. Potentially identifying behavioral examples were not used to illustrate the findings in this paper.

Transcribing and Coding

All interviews were transcribed by the psychology student in a Word document. The mock interview was used for practice. All transcribed interviews were then proof read by the first author (and interviewer). All interviews were uploaded into MAXQDA 2018, a qualitative data analysis software package (Verbi, Germany) that we used for coding and further analysis.

PCL-R Records

For all cases discussed by the interviewees a PCL-R score had been determined at the beginning of treatment, as part of the regular assessment procedure in the clinics involved in this study. Available file information from the PCL-R assessment included criminal records, police records, reports from previous institutes and from prison, and information acquired from relatives, former employers, and schools. Each patient was scored independently by two raters, who then determined a final consensus-score

together. The raters were licensed psychologists who had completed a three-day training in scoring the PCL-R.

In his 1991 manual, Hare recommended a cut score of 30 to classify a patient as psychopathic, based on an American sample. However, it has become evident that in European samples a lower cut score may be more appropriate (Mokros et al., 2013). Based on a meta-analysis of samples from German speaking countries it has been suggested that a score of 25 reflects the same level of psychopathy as a score of 30 in North American samples (Mokros et al., 2013). We instructed the participants to include cases with a score of 28 or higher, to ensure that all cases could be considered highly psychopathic (i.e. well beyond the European cut-off). The standard error of measurement for standard PCL-R assessments is approximately three points (Hare, 2003). See Table 2 for the total scores of the patients discussed by the participants. All patients were male offenders. Successful cases had a mean total score of 33.33 (SD = 2.58); failed cases had a mean total score of 33.56 (SD = 2.85).

Instruments

Table 2 *PCL-R Scores and Age of Patients Discussed in the Interviews*

Year of birth success	PCL-R success	Year of birth failure	PCL-R failure
1968	31.6	1985	32
1979	28.9	1978	31
1974	34.7	1966	30
1979	31.6	1975	34
1983	33.7	1967	34.4
1982	35.6	1984	38.9
1974	34.7	1961	31
1963	37	1986	35.3
1959	34.4	1968	30.6
1964	29.5	1959	36
1967	35	1971	36

Note. The fractions in the total scores (e.g., 31.6) result from the prorating of scores when items have been omitted, in accordance with the instructions in the PCL-R manual (Hare, 2003).

Interview
The semi-structured interview was developed by the first author, and revised after feedback from two of the senior researchers. Subsequently, a mock interview was held with a licensed psychologist of one of the hosting clinics. This led to a final revision; see Table 1. Each of the questions could be followed up with probes for further clarification. As can be seen in Table 1, the interview first focused on patient characteristics, then switched to factors related to the treatment team that may have influenced the course of treatment, and finally provided the opportunity to describe further pertinent factors not yet discussed. Participants were asked to illustrate their answers with specific examples from clinical practice, as illustrative examples often yield more elaborate material than purely theoretical discussion.

PCL-R
The Psychopathy Checklist-Revised (Hare, 1991, 2003) consists of 20 items, which can be scored 0 (*definitely does not apply*), 1 (*may apply or partly applies*), or 2 (*definitely applies*), leading to a possible maximum score of 40. This maximum score is considered to represent a 'prototypical psychopath'. Extensive psychometric properties have been documented in the manual (Hare, 2003). Inter-rater reliability for a sample from the primary hospital involved in this study has been estimated previously, largely based on the same pairs of raters (Hildebrand et al., 2002). The single measure ICC was .88 for the PCL-R total score.

In the first edition of the PCL-R (Hare, 1991), exploratory factor analysis led to the identification of two underlying factors, a model that was replicated several times (Hare & Neumann, 2008). Factor 1 contains the personality traits typically associated with psychopathy; shallow affect, lack of empathy, and a manipulative, arrogant interpersonal style. Factor 2 reflects a chronically impulsive, aggressive and antisocial lifestyle. In the second edition of the PCL-R (Hare, 2003), Hare revised the original model into the more fine-grained Four-Facet model, which was also replicated repeatedly (Neumann et al., 2015). Facet 1 (labeled 'Interpersonal') refers to a glib, arrogant and deceptive interpersonal style; Facet 2 ('Affective') contains items describing the callous and unemotional traits; Facet 3 ('Lifestyle') describes an impulsive and irresponsible lifestyle; Facet 4 ('Antisocial') refers to aggressive and antisocial behavior.

Analytic Strategy

Consensus Coding and Audit
The guidelines explicated in the edited CQR handbook (Hill, 2012) were followed in our data analysis. CQR relies on a consensus process, in which the members of the primary research team first examine the data independently and then meet to discuss their ideas until the entire team agrees on the coding. This process is repeated several times. In this study, it was used first to identify the domains, then to assign all relevant text segments into one or more domains, subsequently to develop categories, and finally to assign all core ideas (summaries of text segments) to the categories. After each step

in the consensus process, the auditor (final author) was consulted for feedback. This led to a relabeling and a subtle reorganization of the domains, and to a collapsing of several categories into a less elaborate but more straightforward system (see Table 3 for the final result).

Table 3 *Results of Cross Analysis*

Domains / Categories	Frequency
Domain 1. Patient characteristics	
- Not measured by PCL-R	general
- Related to PCL-R facet 1 (interpersonal)	general
- Related to PCL-R facet 2 (affective)	general
- Related to PCL-R facet 3 (lifestyle)	typical
- Related to PCL-R facet 4 (antisocial)	variant
Domain 2. Team characteristics	
- Level of expertise	typical
- Sufficient emotional distance yes/no	typical
- Consensus about diagnosis and treatment yes/no	variant
- Continuity in treatment team yes/no	variant
Domain 3. Treatment strategies	
- Underlying treatment philosophy	general
- Reaction to treatment interfering behavior	typical
- Specific interventions	typical
Domain 4. Team-patient interactions	
- Working relationship	variant
- Development of trust	variant
- Splitting	variant
- Fear or aversion	variant
- Punishment	variant
Domain 5. Social Network	
- Prosocial network (yes/no)	typical
- Network support for treatment	typical

Note. A frequency of *general* was coded for appearance of the category in ten or eleven (all) interviews, *typical* for appearance in six to nine interviews, *variant* for two to five interviews.

Domains and Core Ideas

There are two methods for constructing a domain list (Thompson et al., 2012). The first is to start with a list based on themes that emerge from a review of the literature, and to modify this list during coding to ensure that it accurately reflects the data under investigation. Considering the lack of literature in our area of research, this was not a viable option. We therefore used the second method, which involves an inductive approach. We used three randomly chosen interviews to see what topics emerged, and constructed an initial list that was modified throughout the coding and the audit discussions. The next step in the data analysis was to capture the content of each coded text fragment into a so-called 'core idea', a succinct version of what the participant said that is both more clear and concise (Thompson et al., 2012). In this study, the developing of core ideas was not prepared independently by the members of the team, but was done by the primary team immediately upon the coding of the domains.

Cross-Analysis and Categories

Once all relevant text segments of each interview were coded into one or more domains, we started the cross-analysis, which aims to capture common themes *across* the interviews. Each participant of the primary team independently examined all core ideas per domain and developed a list of categories. In a consensus meeting a joint list was created, which was revised again during the process of coding the categories, and subsequently as a result of the audit discussion (e.g., where possible using the same categories for cases that failed and cases that successfully ended treatment, instead of two separate lists of categories). This resulted in a final product.

Results

Over 11 interviews, 369 text fragments were coded; 356 in five domains, 13 (3.5%) in the domain "other". See Table 3 for the results. The five main domains were labelled as: (a) patient characteristics, (b) team characteristics, (c) treatment strategies, (d) team-patient interactions, and (e) social network. Within the domains, several categories were found. With three exceptions, all categories were applicable to both successful and failed cases. For example, having a social network in a successful case versus lacking such in a failed case were considered two sides of the same coin. Both were coded in the category 'Social network'. Therefore, all categories, except 'Splitting', 'Fear or aversion', and 'Punishment', contain both failed and successful cases. Following CQR methodology (Hill et al., 2005), the frequency labels of "general", "typical", and "variant" were assigned to all categories. *General* applied to categories that were found in all interviews, or all but one (ten or eleven interviews). *Typical* applied to categories found in more than half of the interviews (six to nine), and *variant* to those found in two to five interviews. Below, we describe all categories per domain, starting out with the most representative category. To optimally convey its content, we provided illustrative interview quotations of most categories.

Domain 1: Patient Characteristics

As can be seen in Table 3, patient characteristics relevant to success or failure in treatment could be divided into characteristics associated with facets of the PCL-R

and characteristics unrelated to the PCL-R facets. Facet 1 and 2, the Interpersonal and Affective Facets of the PCL-R, and certain characteristics outside the PCL-R were all mentioned quite frequently; i.e. considered *general*. Facet 3 (Impulsive) was mentioned in more than half of the interviews (*typical*), while Lifestyle Facet 4 was only found in a handful of interviews (*variant*).

Not Measured by the PCL-R (General)

This category referred to age, motivation, and comorbidity. Older psychopathic offenders were thought to be more successful relative to younger ones, for two reasons: physical deterioration and having 'calmed down' over the course of life. Admittedly, "older" is an ambiguous term, but when clinicians mentioned a specific number, it was over forty or even fifty years of age. Second, some form of motivation was described as vital to treatment, even if merely extrinsic or exclusively based on a wish to be free from interference by the criminal justice system. See the following excerpt:

> It appears that he came to a point in his life that he thought, "wait.... I am now forty, I am stuck in the TBS, maybe I should try to make something of my life in a different way". Not at all from any intrinsic motivation for treatment, but just simply based on the confrontation with a complete lack of perspective for the future otherwise, and the fact that he had somehow learnt to reflect a little bit on his life. (Treatment success, participant K)

Comorbidity was generally thought to be a drawback, except in the case of features of borderline personality disorder, as defined by the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; DSM-5; American Psychiatric Association, 2013). Several participants felt that having some borderline traits in the emotional domain somehow 'softened' the psychopathic manifestation and made it more susceptible to supportive interventions.

Related to Facet 1: Superficial Charm, Unreliability, Narcissism (General)

This is the Interpersonal Facet of the PCL-R, relating to superficial charm, unreliability and narcissism. The degree to which offenders were able to accept the authority of their treatment providers despite their grandiose sense of self, was described as an important factor in success or failure. Secondly, the degree to which offenders behaved in reliable ways was thought to be crucial. Some lack of reliability was considered normal in this population, but excessive deceit and fickleness during treatment impeded treatment, according to many informants. See the following excerpts:

> The first thing that comes to mind is just the constant deceit, and conning, and externalizing, and saying yes while meaning no. Voicing good intentions and then letting them fail. You can't tell anymore whether it is impulsivity or lack of frustration tolerance (which I do think are important) or just plain deception, agreeing to do something but secretly making his own plans. So, I think conning and manipulation are the most important factors. (Treatment failure, participant G)
> Deception is the most important factor in failure. The TBS-system does not have an answer to that. Many other factors can be treated or managed, like impulsivity

or lack of empathy, but we have not been able to find a creative way to deal with deception. (Treatment failure, participant K)

Related to Facet 2: Shallow Affect, Lack of Empathy, Failure to Take Responsibility (General)
Facet 2 is the Affective Facet of the PCL-R, and is related to shallow affect, lack of empathy and failure to take responsibility. Participants felt that the degree of callousness was an influential factor in treatment. Treatment providers reported that being able to work with a patient required some minimal emotional connection. If, however, the patient was deemed entirely instrumental or even sadistic, treatment often failed. The same counted for a complete failure to take responsibility for the offenses that led to the TBS. See the following excerpts:

He was someone who was not just indifferent with respect to other people's interests, he actually enjoyed crossing boundaries and physically hurting people. (Treatment failure, participant K)

He could be humorous, entertaining. And now and then he would show a glimpse of empathy, for example admitting that he was making life very difficult for us. (Treatment success, participant J)

Related to Facet 3: Impulsivity, Thrill Seeking, Irresponsible Lifestyle (Typical)
The items of Facet 3 are related to impulsivity, thrill seeking, and an irresponsible, parasitic lifestyle. Again, the degree to which these issues were characteristic for the patient was reported as critical to treatment outcome. Reportedly, some degree of impulsivity can be managed, but within limits. See the following excerpts:

He did have some impulsivity, but it was just within limits. He didn't overdo it. [...] Otherwise it would have been a very unfavorable factor. (Treatment success, participant E)

It was only temporarily that he managed to stick to the program. As long as he was kept inside the clinic, for a while it went okay. But as soon as he was given more freedom, his behavior tended to deteriorate. (Treatment failure, participant A)

Related to Facet 4: Aggressive and Antisocial Behavior (Variant)
Facet 4 is related to aggressive and antisocial behavior. Apparently, this is considered less influential with regard to the course of treatment, as the representativeness of this category is *variant*. Early onset of problematic behavior was named as a risk factor for failure of treatment, as well as extreme aggression while admitted to the TBS-hospital:

In my view, what was important is that he was really a very aggressive man. He was also very aggressive inside the clinic. Not all patients have this, or not to that degree. (Treatment failure, participant H)

Domain 2: Team Characteristics

Within this domain, four categories were coded. The first two were labeled *typical*, while the third and the fourth were *variant*.

Level of Expertise (Typical)

A majority of informants emphasized that teams working with psychopathic offenders need to have specialized training. They need to have sufficient knowledge about psychopathy, and the requisite skills, patience, and experience to deal with difficult situations. See the following excerpt:

A team needs to be prepared for unexpected situations. With these patients, you just know that they will occur, and you need to be prepared. They shouldn't surprise you. So, it is essential that a team is knowledgeable about psychopathy. (Treatment success, participant K)

Sufficient Emotional Distance (Typical)

Likewise, a majority of informants indicated that the ability of team members to see the behavior of the patient as the patient's problem or handicap, and not to take it personally, was important. See the following excerpt:

[It is important that] you don't have too many expectations, and that you can really think, "OK, this is part of the job. [...] We knew he was unreliable, but he doesn't do it to spite us, it's just how he is with everyone. Don't take it personally". You try not to mix your own moral values with your work, so to speak. It's quite difficult, and you need your team to help you with this. (Treatment success, participant J)

Consensus about Diagnosis and Treatment (Variant)

Participants reported that disagreement in a team about the goals of treatment and how to interpret the patient's behavior often led to failure, while consensus on the treatment strategy appeared to be a requirement for success. The following excerpt illustrates this:

The team did not really agree on limit setting. [...] For example, when he had used drugs, part of the team felt that his permission for leave outside the clinic needed to be restricted for a long time, while others were of the opinion that this wasn't really necessary. Of course, in such situations a decision was always reached, but you could sense that underneath, the team was divided. (Treatment failure, participant B)

Continuity in Treatment Team (Variant)

Participants reported that too much staff turnaround in teams leads to loss of salient experience in the team and to disruption of the treatment process for the patient. See the following excerpt:

New team members generally have a need to work with clear rules. On the other hand, with psychopathic patients, you need to be able to play a little bit, to know when to be flexible and when to set clear limits to certain behavior. In my opinion, you just have to be very experienced to be able to do this well. The constant turnover of young and inexperienced men and women in our clinic makes this impossible. (Treatment failure, participant D)

Domain 3: Treatment Strategies

Underlying Treatment Philosophy (General)
Together with three of the five categories in the domain patient characteristics, this was the only other *general* (as in almost universally reported) category. Informants reported several elements that they felt were crucial basic treatment principles. The first was creating a clear perspective or goal for the patient. See the following excerpts:

> [At first] it was completely unclear what the route towards rehabilitation should be. And he had no idea what his own goal was. So he was sort of at the mercy of his fellow patients in the group. [...] But then at one point, it became clear in which residence he could live if only he would commit himself to treatment and gain more liberties. At that point he realized, "I would really like that." (Treatment success, participant D)
> The thing is, we didn't really have a good plan [...] for the route towards rehabilitation. We just tried some things. And when that didn't go too badly, we said, "OK, let's try some more", but we never really had a clear goal for this patient. [...] We didn't discuss this openly with him. [...] We gradually and very carefully extended his possibilities for leave outside the clinic, but he didn't understand why we went so slowly. This was an explosive mix. (Treatment failure, participant K)

Informants also felt that a relatively limited inpatient phase worked best, whereas there should be an extended phase outside the clinic, with gradual, stepwise exposure to situations associated with risk for recidivism. Additionally, participants reported that encouraging patients to take responsibility for their own actions, being clear about expectations but allowing some room for mistakes often worked well. See the following excerpt:

> He did have some relapses into drug abuse, but the deal was that if he reported these voluntarily, he would not be readmitted to the clinic, unless it happened very often. Beforehand, he had also been asked to think about what kind of arrangement would help him to stay clean permanently. So, he was emphatically invited to take his own responsibility. And he was also given a lot of autonomy. I think that was good for him, that we didn't take all responsibility away from him, but on the contrary, stimulated him to think for himself. (Treatment success, participant A)

Reaction to Treatment Interfering Behavior (Typical)
Treatment interfering behavior was considered part and parcel to the treatment of psychopathic patients. When this category was referenced, participants were quite unanimous in how best to deal with it: by immediate and predictable limit setting, and if necessary (for security reasons) by imposing a short period of restriction of freedom, followed by resumption of the treatment program. See the following excerpts:

> So, the moment [...] he does not stick to the rules, a direct consequence should follow. A short intervention. And then treatment continues. And I think in this case, with this patient, this happened a lot. At some point, you get a kind of behavioral training. He started to understand the right thing to do, or at least how to stay on course and not get in trouble all the time. (Treatment success, participant I)

[The new team] worked with short and relatively predictable interventions. [...] And also with a proportionate response to what the patient did. [...] For example, when he had done something wrong on his mobile phone. In that case he wasn't forced to hand in the entire phone, but he had to remove the app involved. While before, with the other team, they retained his mobile phone for an indefinite period of time. (Treatment success, participant C)

Specific Interventions (Typical)
Participants *typically* reported various kinds of interventions that had been helpful in specific cases, but no clear picture emerged. This was due to the relatively wide range of interventions, and to the fact that some interventions appeared to have equivocal results. For example, Eye Movement Desensitization and Reprocessing (EMDR) was named as a factor in a successful as well as an unsuccessful case.

Domain 4: Team-Patient Interactions
All five categories in this domain were of *variant* frequency.

Working Relationship (Variant)
This category refers to the situation when, irrespective of the shallow affect and lack of commitment of the patient, to some modest extent, a working relationship still developed with the staff. Interviewees described that some psychopathic patients inspire a certain amount of sympathy; for example, because of a traumatic life history. Possibly patients sensed this sympathy and felt less defensive; for treatment providers it can help in enduring treatment interfering- or other problematic behavior. When all sympathy is lacking, treatment is more likely to fail, according to the participants. See the following excerpt:

I felt there was some degree of connection between us. Of course, with such a high score on the PCL-R you know the connection probably wasn't very strong for him. He did deceive us a lot. But I think he also genuinely felt our empathy for him. (Treatment success, participant G)

Development of Trust (Variant)
Trustworthiness is uncommon in psychopathic offenders. Vice versa, psychopaths are not very likely to trust their therapists. Participants reported that nevertheless, with some patients, a certain amount of mutual trust developed in the course of treatment, or at least a kind of 'reciprocal predictability'. On the other hand, in some of the described failed cases, mutual distrust did not abate at all, and the behavior of the patient remained unpredictable for the team. See the following excerpt:

This was a patient with a history of extremely violent behavior, both outside and inside the clinic. Also, he had managed to have a secret intimate relationship with a staff member. [...] At the same time, he had periods where he seemed calm and in control. But we never felt that we could predict his behavior, there was always a feeling of tension and distrust. Ultimately, this made it impossible to make progress in treatment. (Treatment failure, participant H)

Splitting (Variant)

This category refers to the situation in which a patient, for example through charm or manipulation, created a division in the staff. Members of the team disagreed, or even argued about the patient and what would be the optimal treatment planning, and could not overcome their differences of opinion in an effective manner. This was a precursor to several failures of treatment. See the following excerpt:

> He allowed some members of the team to be closer to him, while he kept other team members more at a distance. The result was that those whom he appeared to trust more, were more positive about the treatment effects and more willing to make progress, while others remained more suspicious of him. We did not manage to solve this. (Treatment failure, participant B)

Fear or Aversion (Variant)

Sometimes patients were so hostile or dangerous even within the clinic (i.e., causing aggressive incidents), or had committed such horrific crimes, that team members experienced fear and/or aversion. See the following excerpt with an example of aversion:

> This patient was part of a group that met in some foreign country to participate in extreme sexual abuse of children. The more you paid, the more you were allowed to do. It was really very horrible to read his file. And also, to listen to him minimizing the impact of his behavior. [...] It was very difficult to see this offender as a patient, and to contain the negative countertransference. (Treatment failure, participant G)

Punishment (Variant)

In some cases, teams did not find a way to tolerate and manage problem behavior (such as breaking institutional rules or violating conditions for leave outside the hospital; e.g., using drugs, smuggling a smart phone into the facility), leading to ineffective punishment of the patient instead of effective limit setting (e.g., restrictions in the freedom of movement in or outside the clinic, which the respondents felt were disproportionate). See the following excerpt:

> When he broke the rules, I felt the staff reacted with anger, punitively, as if they felt betrayed. And they tried to appeal to his conscience, I think. But it didn't work out that way. What happened was that he was more and more convinced that he could never do it right. (Treatment failure, participant B)

Domain 5: Social Network

Both categories in this domain had a frequency of *typical*.

Prosocial Network (Typical)

A majority of participants reported that simply having a social network that was not overly antisocial was a positive factor; e.g. through offering support and improving motivation. Conversely, a complete absence of social contacts, or mainly antisocial connections, did not bode well. See the following excerpt:

> He reconnected with his old antisocial friends, but they now all had families of their own. They had become more normal. They were still involved in, for example,

small scale, illegal trade, but that was it. No more violence. (Treatment success, participant H)

Network Support for Treatment (Typical)

Both active support of, and opposition to the TBS order and the treatment were important factors. See the following excerpts:

[The family felt that] the TBS order was completely unjustified. Nothing wrong with him. They thought we were ridiculous. What did we think we were doing? (Treatment failure, participant F)

His social network was very important. His mother, for example, told us he was a different man when he used alcohol. She made sure that during family visits, no alcohol was available to anyone. The patient himself thought he could manage a beer now and then, but because of his mother's attitude he never tried. This was very important in his treatment. (Treatment success, participant B).

Discussion

In this study, we used a qualitative research design to explore with expert clinicians factors that in their view influenced successful completion versus drop-out in the forensic psychiatric treatment of highly psychopathic offenders; individual patient characteristics as well as characteristics of the treatment or team, were sought. A variety of factors emerged. First, a number of general forensic treatment-impeding factors were mentioned as equally relevant for offenders with severe psychopathic traits: age, motivation, and patients' network. While it may seem unsurprising that factors relevant for forensic treatment in general also appear relevant in the treatment of psychopathic offenders, some of these observations are still notably at odds with general beliefs about psychopathy. A good example concerns the age of the offender. In general, it is believed that men high on psychopathic traits start their offending behavior at a young age and, in contrast to 'ordinary' antisocial offenders, do not easily quiet down. However, according to our informants, there still appears to be an age at which they do quiet down. In similar vein, the issue of a minimal amount of some kind of motivation is remarkable. It was repeatedly emphasized that it did not matter how 'thin' or self-centered the motivation might be; any and all forms of motivation were deemed helpful and some minimum essential. In other words: Psychopathic patients don't have to do it for the *right* reasons, as long as they can find any reason. Another interesting finding is that the network, in the eyes of the interviewees, does not have to be exemplary prosocial, just not too antisocial. A little rule violation is not a problem per se, as long as there are clear limits within the network as to the unacceptability of violence. We know of no studies that examine the role of involving family and peers in the treatment of adult psychopathic offenders specifically.

Several factors were related to the treatment and the team. Relevant treatment characteristics for psychopathic patients mostly referred to an underlying philosophy or framework. This framework included (a) matching clear goals and perspectives for the patient with clear and concrete expectations from the patient, and (b) keeping the

inpatient treatment relatively short, followed by a prolonged and gradual outpatient trajectory. Other important suggestions our interviewees made were tolerating some rule violation, choosing your battles, and applying clear, predictable punishment of short duration, where necessary. This seems applicable to the treatment of all forensic patients, but likely needs more emphasis for psychopathic patients, who are of course among the hardest to treat within this group.

Team characteristics were regularly mentioned as relevant. However, most of the characteristics are relevant factors for any treatment, such as consensus on diagnosis and treatment trajectory. A team characteristic related to psychopathy was expertise on psychopathy in general and, more specific, the ability to keep sufficient emotional distance. It appears that a team that is well informed on the nature and characteristics of psychopathy is therefore better able to view psychopathy as a deficiency and not feel personally betrayed by the failures of their patients. Continuity of treatment providers was also deemed important. This finding seems somewhat counterintuitive. Why would this matter for psychopaths who are believed to be notoriously uninterested in and incapable of attachment? The overall findings of this study seem to indicate that the prospects of successful treatment for psychopaths are better if any relationship, no matter how unstable, pragmatic/opportunistic or superficial, can still grow. However, it seems that these relationships with highly psychopathic offenders a) take long to develop, and b) easily get damaged. Discontinuity in a team is thought to be one of the disturbing factors.

A number of factors more specific to psychopathy emerged from the study as relevant. Patient characteristics associated with Facets 1 and 2 of the PCL-R were found in (nearly) all interviews; those associated with Facet 3 in more than half. Facet 4 was considered less influential. This is in line with three of the four studies cited in the introduction (Olver & Wong, 2011; Jeandarme, et al., 2017; Sewall & Olver, 2019). Of note, for all facets, the informants reported that it was the *degree* to which a certain facet was present that counted. That is, some degree of deception, arrogance, callousness, impulsivity and aggression were all to be expected behaviors of these patients during treatment, but only very high levels were thought to be seriously treatment interfering.

The problems arising from overly pronounced Facet 1 or Facet 2 traits were described in two variations: either no relationship or form of cooperation could develop at all (emphasis Facet 2), or a relationship appears to develop, but is destroyed by dishonest and untrustworthy behavior (emphasis Facet 1). It remains unclear whether this dishonest behavior should be seen as planned manipulation or that psychopathic patients are fundamentally unable to acknowledge their faults and simply declare their own desired truth, which exists separate from reality. Very high traits on Facet 3, and very high traits versus more moderate traits on Facet 4 were mentioned to a lesser extent and mostly in the form of excessive impulsivity or aggression. Most of the Facet 4 items (Poor Behavioral Controls, Early Behavioral Problems, Juvenile Delinquency, Revocation of Conditional Release, Criminal Versatility) refer to historical facts and behavior, and to a far

lesser extent to personality traits. Therefore, it is not surprising that Facet 4 is less often nominated as a treatment interfering patient characteristic. Nevertheless, Facet 4 was mentioned as relevant in a few cases, specifically regarding patients that caused extremely serious violent incidents within the walls of the facility. Serious violent incidents within the facility caused by psychopathic patients are relatively rare in our experience, but cause strong backlash, and are often the prelude to transfer, according to our informants.

The recurring suggestion that comorbid borderline personality traits may be favorable for treatability of psychopathic patients, is intriguing. How may comorbid borderline traits make psychopathic patients more susceptible to treatment? Stipulating recommendations for further research on psychopathy, Lilienfeld (2018) argued that it is important to incorporate measures of other disorders, especially antisocial, histrionic, borderline, and narcissistic personality disorders, also known as the cluster B personality disorders. He noted that although comorbidity generally is considered to increase impairment, there may be notable exceptions. He cited two studies (Walker et al., 1991; Short et al., 2016) that found that in children and adolescents with conduct disorder, comorbid symptoms of depression or anxiety tended to be associated with better outcome. Lilienfeld hypothesized that the presence of an internalizing disorder may attenuate risk. A similar mechanism may hold for borderline traits comorbid with psychopathy. One might conjecture that presence of emotional borderline features is incompatible with extreme emotional callousness - a factor that was frequently mentioned as a risk factor for treatment failure. We know of no studies that examined the effect of borderline traits on the treatment of psychopathic offenders, and further research seems warranted.

The current study has several limitations. First, we asked clinicians to choose two specific cases, in the expectation that illustrative examples would allow for in-depth discussion. This requirement also ensured that truly psychopathic patients were chosen. However, it is possible that this approach led to clinicians choosing particularly memorable (instead of representative) cases. Another strategy would have been to frame the questions in a more general way, focusing on psychopathic patients in general. At the end of the interview, we did provide the clinicians the opportunity to add general comments (see Table 1; "According to you, are there other factors not yet discussed on the basis of these two cases, which may be essential to take into account when treating psychopathic offenders?"). Second, we did not specify in great detail how to prepare for the case interviews. It became clear during the interviews that some participants had taken the time to study the files, while others appeared to rely predominantly on their memory. It is not immediately clear how this variation may have affected the findings. It bears mentioning that we consistently requested informants to bolster their comments by providing concrete, specific examples to best illustrate their answers. Also, the argument could be made that the most salient factors would have remained in long-term memory anyway. Third, this study focused exclusively on factors related to treatment completion in highly psychopathic patients. While limiting drop-out is an important goal in developing effective treatment for these patients, this study did not focus on the effects of treatment

on reducing recidivism. Additionally, our research design precludes determination of the extent to which these factors of treatment success (or failure) are specific to psychopathic patients. When possible, however, we provided some informed conjectures on this issue. Fourth, no specific interventions for reducing the responsivity problems in psychopathy were consistently identified. Respondents sometimes mentioned a certain form of treatment that was effective (or not) in a specific case, but no general pattern emerged. At this stage, it seems advisable to try various interventions and to not give up too easily. Even if no specialized interventions have yet been developed for psychopathic patients in general, some may still be effective in particular cases. Much more research is needed in this domain. Fifth, for the purpose of homogeneity, all but one of the interviewees were working in the same hospital, albeit in different units. Therefore, the results of this study may be vulnerable to embedded culture effects of this particular setting. Future studies should widen the scope, either by replicating the current study in another hospital, or by including more than one institute in the research design. A final limitation is that this study is limited to the perspective of the treatment providers. As of yet, there are no data on how patients reflect on their experiences in forensic treatment. We argue this remains a worthwhile empirical question. Eliciting the patients' perspective may well be a profitable avenue for the further development of specialized treatment programs. Notwithstanding these limitations, tentative conclusions regarding the ideas of experienced treatment providers about the treatability of psychopathic forensic patients can be drawn. Overall, the results imply that psychopathic patients may be retained in treatment successfully lest they are not exceedingly psychopathic, especially with regards to arrogant and deceitful interpersonal style (Facet 1) and defective affective experience (Facet 2). However, not only specific psychopathic traits are thought to be important, other patient characteristics also appear influential. Older patients and those with a somewhat prosocial and supportive network may fare better. If some form of a working relationship can develop, no matter how shallow, and some form of motivation can be found, no matter how self-centered, treatment of psychopathic patients can proceed successfully. Future research into the treatability of psychopathic patients should therefore take into account other patient characteristics than psychopathic traits only.

Of note, our study also indicates that certain characteristics of the treatment framework were deemed highly relevant for treatment success or failure. A treatment program that stipulates clear and concrete goals and expectations, provides a long and gradual resocialization trajectory, and is offered by a knowledgeable, cohesive, and stable team, appears to be more successful. This means that forensic treatment as usual may need to be adapted for psychopathic patients. For example, not all forensic patients need the extensive outpatient phase of treatment that is recommended in our study. Additionally, teams working with psychopathic patients probably need extra training, not only to limit the possible detrimental effects of working with these patients, but also to enhance their ability to retain these patients in treatment in the first place. All these suggestions of course need to be validated in future (quantitative) studies. At this point, developing a treatment framework using some of these suggestions appears to be a priority. In this respect, it is intriguing to recall that 25 years ago, after many years of clinical

pessimism about the prognosis, Dialectical Behavior Therapy (DBT; Linehan, 1993) was the first evidence-based psychotherapy for borderline personality disorder. One of the explicit goals of DBT was limiting attrition, and the first clinical trial showed significantly less drop-out when compared to treatment as usual, 16.7% versus 50% (Linehan et al., 1991). In her book, Linehan first devoted no less than 200 pages to describing the theoretical framework and the specification of the parameters for treatment, including integration of supervision and consultation for therapists, before introducing any concrete therapeutic techniques for patients. As the treatment of psychopathic patients is without a doubt just as challenging for treatment providers, developing such a manual would be of great value for this clinical condition as well. Some of the topics emerging from this study may eventually be essential reading in the first chapters.

References

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Author.

Burkard, A. W., Knox, S., & Hill, C. E. (2012). Data collection. In C. E. Hill (Ed.), *Consensual qualitative research: A practical resource for investigating social science phenomena* (pp. 83-101). American Psychological Association.

Cullen, A. E., Soria, C., Clarke, A. Y., Dean, K., & Fahy, T. (2011). Factors predicting dropout from the reasoning and rehabilitation program with mentally disordered offenders. *Criminal Justice and Behavior, 38*(3), 217-230. https://doi.org/10.1177/0093854810393659

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

Hare, R. D. (2003). *Hare Psychopathy Checklist-Revised (PCL-R): 2nd edition.* 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

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

Hecht, L. K., Latzman, R. D., & Lilienfeld, S. O. (2018). *The psychological treatment of psychopathy: Theory and research.* In D. David, S. J. Lynn, & G. H.Montgomery (Eds.), *Evidence-based psychotherapy: The state of the science and practice* (p. 271-298). Wiley-Blackwell.

Hildebrand, M., de Ruiter, C., de Vogel, V., & van der Wolf, P. (2002). Reliability and factor structure of the Dutch language version of Hare's Psychopathy Checklist-Revised. *International Journal of Forensic Mental Health, 1*(2), 139-154. https://doi.org/10.1080/14999013.2002.10471169

Hill, C. E. (Ed.). (2012). *Consensual qualitative research: A practical resource for investigating social science phenomena.* American Psychological Association.

Hill, C. E., Knox, S., Thompson, B. J., Williams, E. N., Hess, S. A., & Ladany, N. (2005). Consensual qualitative research: An update. *Journal of Counseling Psychology, 52*(2), 196-205. https://doi.org/10.1037/0022-0167.52.2.196

Hill, C. E., Thompson, B. J., & Williams, E. N. (1997). A guide to conducting consensual qualitative research. *The Counseling Psychologist, 25*(4), 517-572. https://doi.org/10.1177/0011000097254001

Jeandarme, I., Pouls, C., Oei, T. I., & Bogaerts, S. (2017). Forensic psychiatric patients with comorbid psychopathy: Double trouble? *International Journal of Forensic Mental Health, 16*(2), 149-160. https://doi.org/10.1080/14999013.2017.1286414

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

Kröger, U., Beek, D. van, Wolf, P van der, Klein Haneveld, E., Geest, H. van & Geraerts, R. (2014). *Treatment of psychopathy: A mission impossible?* Forum Educatief.

Ladany, N.,Thompson, B. J., & Hill, C. E. (2012). Cross-Analysis. In C. E. Hill (Ed.), *Consensual qualitative research: A practical resource for investigating social science phenomena* (pp. 117-134). American Psychological Association.

Lilienfeld, S. O. (2018). The multidimensional nature of psychopathy: Five recommendations for research. *Journal of Psychopathology and Behavioral Assessment, 40*(1), 79-85. https://doi.org/10.1007/s10862-018-9657-7

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder.* The Guilford Press.

Linehan, M. M., Armstrong, H. E., Suarez, A., Allmon, D., & Heard, H. L. (1991). Cognitive-behavioral treatment of chronically parasuicidal borderline patients. *Archives of General Psychiatry, 48*(12), 1060-1064. https://doi.org/10.1001/archpsyc.1991.01810360024003

Mokros, A., Hollerbach, P., Vohs, K., Nitschke, J., Eher, R., & Habermeyer, E. (2013). Normative data for the psychopathy checklist–revised in German-speaking countries: A meta-analysis. *Criminal Justice and Behavior, 40*(12), 1397-1412. https://doi.org/10.1177/0093854813492519

Neumann, C. S., Hare, R. D., & Pardini, D. A. (2015). Antisociality and the construct of psychopathy: Data from across the globe. *Journal of personality, 83*(6), 678-692. https://doi.org/10.1111/jopy.12127

Ogloff, J. R., Wong, S., & Greenwood, A. (1990). Treating criminal psychopaths in a therapeutic community program. *Behavioral Sciences & the Law, 8*(2), 181-190. https://doi.org/10.1002/bsl.2370080210

Olver, M. E., & Wong, S. C. (2009). Therapeutic responses of psychopathic sexual offenders: Treatment attrition, therapeutic change, and long-term recidivism. *Journal of Consulting and Clinical Psychology, 77*(2), 328-336. https://doi.org/10.1037/a0015001

Olver, M. E., & Wong, S. (2011). Predictors of sex offender treatment dropout: Psychopathy, sex offender risk, and responsivity implications. *Psychology, Crime & Law, 17*(5), 457-471. https://doi.org/10.1080/10683160903318876

Polaschek, D. L. L. (2014). Adult criminals with psychopathy: Common beliefs about treatability and change have little empirical support. *Current Directions in Psychological Science, 23*(4), 296-301. https://doi.org/10.1177/0963721414535211

Polaschek, D. L. L., & Skeem, J. L. (2018). Treatment of adults and juveniles with psychopathy. In C. Patrick (Ed.), *Handbook of psychopathy (2nd ed.)* (pp. 710-731). The Guilford Press.

Reidy, D. E., Kearns, M. C., & DeGue, S. (2013). Reducing psychopathic violence: A review of the treatment literature. *Agression and Violent Behavior, 18*(5), 527-538. https://doi.org/10.1016/j.avb.2013.07.008

Sewal, 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

Short, R. M., Sonuga-Barke, E. J., Adams, W. J., & Fairchild, G. (2016). Does comorbid anxiety counteract emotion recognition deficits in conduct disorder? *Journal of Child Psychology and Psychiatry, 57*(8), 917-926. https://doi.org/10.1111/jcpp.12544

Sim, W., Huang, T. C., & Hill, C. E. (2012). Biases and expectations. In C. E. Hill (Ed.), *Consensual qualitative research: A practical resource for investigating social science phenomena* (pp. 59-69). American Psychological Association. Tbs inzichtelijk (2019, April). https://www.google.com/url?sa=t&rct=j&q=&esrc= s&source=web&cd=&ved=2ahUKEwju_ZHc_0P1AhU1gv0HHR- 3Bf0QFn0ECAQQAQ&url=https%3A%2F%2Fwww.forensischezorg.nl%2Ffiles% 2Finfographic_tbs_april_2019.pdf&usg=AOvVaw24qBGS1Vud0nye0MJ0DFz7

Thompson, B. J., Vivino, B. L., & Hill, C. E. (2012). Coding the data: Domains and core ideas. In C. E. Hill (Ed.), *Consensual qualitative research: A practical resource for investigating social science phenomena* (pp. 103-116). American Psychological Association.

Vertommen, H., Verheul, R., Ruiter, C. de, & Hildebrand, M. (2002). *Hare's Psychopathie Checklist: Handleiding [Hare Psychopathy Checklist-Revised]*. Swets and Zeitlinger.

Walker, J. L., Lahey, B. B., Russo, M. F., Frick, P. J., Christ, M. A. G., McBurnett, K., Loeber, R., Stouthamer-Loeber, M., & Green, S. M. (1991). Anxiety, inhibition, and conduct disorder in children: Relations to social impairment. *Journal of the American Academy of Child and Adolescent Psychiatry, 30*(2), 187-191. https://doi.org/10.1097/00004583-199103000-00004

Wong, S. C. P., & Hare, R. D. (2005). *Guidelines for a psychopathy treatment program.* Multi-Health Systems.

Using MMPI-2-RF Correlates to Elucidate the PCL-R and its Four Facets in a Sample of Male Forensic Psychiatric Patients

Abstract

This study documents the associations between the Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008) scale scores and the Psychopathy Checklist Revised (PCL-R; Hare, 2003) facet scores in a forensic psychiatric sample. Objectives were to determine how the MMPI-2-RF scales may enhance substantive understanding of the nature of the four PCL-R facets and to discern possible implications for the treatment of psychopathic patients. A sample of 127 male forensic psychiatric offenders admitted to a Dutch forensic psychiatric hospital completed the PCL-R and the MMPI-2. Exploratory stepwise regression analyses assessed the prediction of the PCL-R total and its facet scores from MMPI-2-RF scales at its three hierarchical levels. Conceptually meaningful results emerged at each level of the MMPI-2-RF hierarchy, including several consistent differences between predictor sets across the facets. Interestingly, ideas of persecution (RC6) was a specific predictor of PCL-R Facet 2, a facet noted for its association with treatment failure. Results are compared and contrasted to the extant body of empirical work to date, and some tentative clinical implications are offered.

In best practice forensic settings, psychopathy is frequently assessed by administration of the Hare Psychopathy Check List-Revised (PCL-R; Hare, 1991, 2003). Currently, the PCL-R operationalizes the construct of psychopathy via four specific, correlated facets. In the field of psychopathy this is a leading contemporary model, albeit not the only one. The MMPI family of instruments operationalizes a predominant model of psychopathology and personality. Its most recent version, the Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008), holds particular heuristic promise, as it boasts a multi-level, hierarchical set up that is more in line with modern theorizing about the nature of psychopathology. We held that the MMPI-2-RF and PCL-R models might amplify and inform each other, especially when the PCL-R is examined from a more fine-grained perspective distinguishing four facets, rather than the traditional two factors. The present study aims to map the MMPI-2-RF onto the Four-Facet model of the PCL-R.

The PCL-R (Hare, 1991, 2003) was designed to reliably measure the clinical construct of psychopathy. Its maximum score (of 40) is considered to represent the "prototypical psychopath". Using exploratory factor analysis with the data set in the first edition of the PCL-R (Hare, 1991), 17 of the 20 items were originally divided into two factors, with three items loading on neither factor. This Two-Factor model was replicated several times using confirmatory factor analysis (Hare & Neumann, 2008). Factor 1 contained the personality traits typically associated with psychopathy; these reflect shallow affect and a manipulative, remorseless, and arrogant interpersonal style. Factor 2 reflected a chronically unstable, aggressive, and antisocial lifestyle. Cooke and Michie (2001) were unable to replicate the Two-Factor model in their study, using confirmatory factor analysis, and developed and cross-validated a hierarchical Three-Factor model. In this model, the concept of psychopathy is underpinned by the following three factors: an arrogant and deceitful interpersonal style, deficient affective experience, and impulsive and irresponsible behavioral style. Their model implies that criminality is not a core feature of psychopathy but rather the consequence of the three core factors. More recently, and largely in response to Cooke and Michie's Three-Factor model, Hare revised his original model (Hare, 2003) to include four facets (called *facets* to distinguish them from the original two factors in name as well as in number). This new Four-Facet model augments the Cooke and Michie's Three-Factor model with a fourth, so-called antisocial facet, using the remaining items from the two-factor model that Cooke and Michie had excluded. The resulting four facets were labeled: Interpersonal (Facet 1), Affective (Facet 2), Lifestyle (Facet 3) and Antisocial (Facet 4). Facet 1 describes a glib, arrogant and deceptive interpersonal style, Facet 2 refers to shallow emotions and lack of empathy, Facet 3 refers to an impulsive and irresponsible lifestyle, and Facet 4 indicates a tendency to violate rules and social norms with aggressive and antisocial behavior (Hare & Neumann, 2008; Neumann et al., 2015). These four facets have been replicated in several international samples (Hare, Neumann & Mokros, in press).

There is ongoing debate whether these four facets should be seen as first-order factors that together form the superordinate factor of psychopathy (Hare et al., in press) or

whether the first three facets define the core psychopathy, with the antisocial facet reflecting merely the behavioral consequences of psychopathy (Cooke & Michie, 2001; Skeem & Cooke, 2010). Framed differently, the debate centers on the question whether antisociality is an essential feature of psychopathy or not. In a recent special issue of the *Journal of Personality*, Miller and Lynam (2014) argued that regardless of whether antisocial behavior should be explicitly part of the assessment of psychopathy, virtually all authors appear to agree that it is at least intimately related to psychopathy. Moreover, from a developmental perspective, it has been shown that early antisocial features predict the development of other features of psychopathy at a later stage (Forsman et al., 2010). Recent psychophysiological research showed that baseline oxytocin levels in high-risk offenders were strongly and specifically related to facet 4 of the PCL-R, and in particular to the items 'early behavioral problems' and 'juvenile delinquency' (Mitchell et al., 2013). These studies provide suggestive evidence that antisocial behavior is more than a mere readout of core personality features. In this paper we will therefore refer to the Four-Facet model.

We see (at least) two principal reasons for selecting the MMPI-2-RF to inform psychopathy. First, in contrast to the PCL-R, which demands the availability and expert evaluation of extensive file information, the instrument can easily be administered. Perhaps as a result, as noted by Archer et al., and Handel (2006), the MMPI-2 (from which the MMPI-2-RF can be derived) is widely used in clinical and forensic settings. Second, juxtaposing the PCL-R facets with a more encompassing model of personality and psychopathology (i.e. the MMPI-2-RF) may help elucidate how the facets are linked to external correlates. For example, early authors on psychopathy (e.g. Karpman, 1946) have speculated that certain subtypes may be more amenable to treatment than others. Only very recently attempts have been made to examine this hypothesis empirically. An important finding was that Facet 2 was associated with treatment drop-out (Olver & Wong, 2011), and uniquely predicted less favorable therapeutic outcomes (Olver et al., 2013). The mechanisms underlying these associations remain unclear and juxtaposing MMPI-2-RF psychopathology indicators with Facet 2 may yield valuable hypotheses. More generally, research has shown that the psychopathy facets are differentially related to external correlates (Neumann & Pardini, 2014), and that individuals with the same PCL-R total score may have distinctive constellations of facet scores. Accordingly, it seems plausible that individuals with psychopathic traits form a heterogeneous group with varying clinical needs, which may be illuminated by the distinctive patterns of MMPI-2-RF scale elevations.

The MMPI-2-Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008; Tellegen & Ben-Porath, 2008) constitutes a logical extension of the development of the MMPI-2 Restructured Clinical (RC) scales (Tellegen et al., 2003). The MMPI-2-RF is comprised of a hierarchical set of scale sets, including the Higher-Order, Restructured Clinical, and Specific Problems and Interest scales. In general, the MMPI-2-RF is designed to provide a range of interpretative possibilities, from a rather broad band approach to personality assessment (i.e., the Higher-Order scales) to a more focused, narrow band level (i.e., the Specific Problems scales).

To date, eight previous studies provide evidence on (expected) associations between the MMPI-2-RF and psychopathy (Anderson et al., 2015; Philips et al., 2013; Sellbom, 2011; Sellbom et al., 2005; Sellbom et al., 2007; Sellbom et al., 2012; Sellbom et al., 2015; Wygant & Sellbom, 2012). The majority of these studies report on associations between the MMPI-2 or the MMPI-2-RF and other self-report measures, predominantly the Psychopathic Personality Inventory (PPI; Lilienfeld & Andrews, 1996). The PPI is comprised of two subscales (named PPI-I Fearless-Dominance and PPI-II Impulsive-Antisociality), which are conceptually similar to the Facets 1/2, and Facets 3/4 of the PCL-R, respectively (e.g., Benning et al., 2003; Patrick et al., 2006). Two studies examined associations between the Psychopathy Checklist Screening Version (PCL-SV; Hart et al., 1995) and a selection of scales from the MMPI-2 (Sellbom et al., 2007) and the MMPI-2-RF (Wygant & Sellbom, 2012), especially the Personality Psychopathology Five scales, PSY-5, and a priori conceptually related scales. To the best of our knowledge, no studies used the full PCL-R, nor its Four-Facet representation. Nevertheless, some tentative conclusions can be gleaned from an inspection of the currently available findings. First, robust (positive) associations have been found across studies between Antisocial Behavior (RC4), Hypomanic Activation (RC9) and the various indices of psychopathy. Mixed evidence has been observed for negative associations with Dysfunctional Negative Emotions (RC7) and Low Positive Emotions (RC2); these associations appeared to be more pronounced when using self-report criteria of psychopathy (e.g., Phillips et al., 2013) than with the PCL:SV (Sellbom et al., 2012). Second, from the perspective of a model of personality disorder-related psychopathology (i.e., PSY-5 scales), robust associations were found for positive associations between lack of inhibition/disconstraint (Disconstraint-revised, DISC-r) and (instrumental) aggressiveness (as measured by Aggressiveness-revised, AGGR-r) and psychopathy indices. Again, mixed evidence was found for the more internalizing personality psychopathology PSY-5 scales (i.e., Negative Emotionality, NEGE-r, and Introversion/ Low Positive Emotionality, INTR-r). With regard to the MMPI-2-RF Specific Problem Scales, the general pattern was that the scales that measure vulnerability, internalizing problems, or inhibitory interpersonal dimensions (including Anxiety, Multiple Specific Fears, Shyness and Social Avoidance; AXY, MSF, SHY, SAV, respectively), correlated negatively with personality characteristics related to psychopathy (especially the Fearless-Dominance factor of the PPI). MMPI-2-RF Specific Problems Scales that assess externalizing problems (like Juvenile Conduct Problems, Aggression, and Substance Abuse; JCP, AGG, SUB, respectively) were generally positively correlated with impulsive and antisocial behaviors related to psychopathy. However, it should be noted that the eight studies are quite heterogeneous in terms of sample size (N ranging from 78 to well over 40,000), composition (i.e., women or men only versus both genders; prison inmates, psychiatric patients, college students, or combinations thereof), and analytic procedures, and differ in their selection of potential MMPI-2/MMPI-2-RF predictors as well as outcome measures.

Given this observed heterogeneity in studies and findings, as well as the consideration that the present study is the first to test the predictive potency of the MMPI-2-RF sets of scales with the full, Four-Facet representation of the PCL-R in a sample of all male

forensic psychiatric patients, we opted for an exploratory analytic strategy testing each full set of MMPI-2-RF scales. However, based on the literature reviewed above, we formulated the following tentative hypotheses. At the RC scale level, we expected, on the one hand, positive associations between RC4/RC9 and the PCL-R total score and its Lifestyle and Antisocial Facets (i.e. Facet 3 and 4); and, on the other hand, negative associations between RC7 and the Interpersonal and Affective Facets (i.e. Facet 1 and 2). With respect to the MMPI-2-RF Specific Problems scales, we expected positive associations between scales that assess externalizing problems (JCP, AGG, and SUB) and the impulsive and antisocial psychopathy facets. Negative associations were expected between one or more of the MMPI-2-RF Specific Problem scales involving fear and anxiety (STW, AXY, BRF, MSF) and the affective features of psychopathy; and between shyness (indexed by SHY) and the Interpersonal PCL-R Facet. Finally, with regard to the PSY-5 scales, we expected positive associations between Disconstraint (DISC-r) / Agressiveness (AGGR-r) and the PCL-R Lifestyle and Antisocial Facets, as well as negative associations between Negative Emotionality (NEGE-r) and the PCL-R Affective and Interpersonal Facets, respectively.

Method

Participants and Procedure
For this study, participants included a total of 139 male patients admitted to a forensic psychiatric hospital in the Netherlands between 1997 and 2009 with a TBS-order. TBS (*ter beschikking stelling*) is mandatory intensive inpatient treatment for high-risk offenders that can be ordered by the Dutch courts as part of a sentence for violent or sexual offences (i.e., assault, manslaughter, murder, rape, child molestation, etc.). All included patients were convicted of a violent offense: 80 (57.6%) committed a life-offense, (i.e., (attempted) murder or manslaughter), 28 (20.1%) committed violent assault, and 31 (22.3%) committed rape/sexual assault against an adult. The goal of treatment is to minimize the risk of reoffending while working towards gradual rehabilitation. Patients are admitted immediately after completing a prison sentence and, as a standard procedure, participate in extensive psychological assessment during the first three months of treatment. This includes, among other instruments, the MMPI-2, the PCL-R, and the Structured Interview for DSM-IV Personality (SIDP-IV; Pfohl et al., 1997). Participants completed the Dutch paper/pencil version of the MMPI-2, from which the MMPI-2-RF scores were derived, and scored according to Dutch norms. As the items contained within the MMPI-2-RF are represented in the larger MMPI-2 item pool, it is possible to extract and score MMPI-2-RF scales from MMPI-2 protocols. According to Tellegen and Ben-Porath (2008), scoring the MMPI-2-RF from the MMPI-2 should not impact the reliability of scale scores or relations with criterion measures. This was confirmed for the Dutch version by Van der Heijden et al. (2010).

Participants' MMPI-2-RF results were excluded from analysis if they produced an invalid profile. For the current study, an invalid MMPI-2-RF was defined as having either a Cannot Say (?) raw score greater than or equal to 18, a True Response Inconsistency (TRIN-r) or Variable Response Inconsistency (VRIN-r) T score greater than or equal to

80, an Infrequent Responses (F-r) T score greater than or equal to 120, an Infrequent Psychopathology Responses (Fp-r) T score greater than or equal to 100, an Uncommon Virtues (L-r) T score greater than or equal to 80, and/or an Adjustment Validity (K-r) T score greater than or equal to 70. These cut scores were derived from recommendations by the authors of the MMPI-2-RF for use with clinical populations (Ben-Porath & Tellegen, 2008), as well as from an examination of the current data. Using these criteria, a total of 12 (8.6%) participants were excluded for producing invalid MMPI-2-RF profiles. Two specifically trained and licensed psychologists independently assessed PCL-R scores, and subsequently met to decide on consensus scoring.

The final group consisted of 127 men. Participants' ages ranged from 19 to 58 (*M* = 32.40, *SD* = 8.49). Ethnic constellation was 70.1% Dutch, 17.2% Afro Caribbean, 5.6% Moroccan, 4.7% Turkish, 2.4% other. Of all participants 4.8% had no education, 49.6% completed primary school but did not finish secondary school/high school, 26.4% completed vocational training, 10.4% completed secondary school/high school, 4.0% had higher degrees, and for 4.8% information about education was missing. Table 1 shows the distribution of personality disorders among the participants with a valid MMPI-2-RF profile. As can be seen, the most prevalent personality disorder (PD) was Antisocial PD (56.7%), followed by Narcissistic PD (42.5%); more than half of the sample met criteria for more than one PD (55.1%).

Instruments

MMPI-2-RF
The Dutch version of the MMPI-2 was administered (Derksen et al., 1993), from which the MMPI-2-RF scales were derived. As described previously, the MMPI-2-RF is a 338 True or False item self-report inventory that assesses an individual's characteristics across multiple domains (e.g., personality, psychopathology, and social/ behavioral

Table 1 *Axis II Diagnosis for Violent Offenders (N=127)*

Axis II diagnosis		Violent offenders
Cluster A	Paranoid	25 (19.7%)
	Schizoid	5 (3.9%)
	Schizotypal	7 (5.5%)
Cluster B	Antisocial	72 (56.7%)
	Borderline	33 (26.0%)
	Histrionic	5 (3.9%)
	Narcissistic	54 (42.5%)
Cluster C	Avoidant	12 (9.4%)
	Dependent	5 (3.9%)
	Obsessive-Compulsive	13 (10.2%)
More than one axis II		70 (55.1%)
No axis II		13 (10.2%)

functioning) and has a number of scales designed to detect potential invalid styles of responding (i.e., non-content or content-based responding). The MMPI-2-RF is comprised of a hierarchical set of scale sets, including the Higher-Order, Restructured Clinical, and Specific Problems Scales and Interest scales. The Higher-Order scales are designed to reflect the structural dimensions underlying the nine Restructured Clinical scales, and factor analytic techniques yielded three distinct dimensions, Emotional/Internalizing Dysfunction (EID), Thought Dysfunction (THD), and Behavioral/Externalizing Dysfunction (BXD). An additional set of 25 scales, the Specific Problems and Interest scales were derived from the full item pool of the original MMPI-2, and are intended to facilitate as well as augment interpretation of both the Higher-Order and Restructured Clinical scales. The Specific Problems scales are arranged into four sets, Somatic/Cognitive, Internalizing, Externalizing, and Interpersonal Problems, each of which contains multiple substantive scales. Finally, the MMPI-2-RF also includes the Personality Psychopathology Five (PSY-5). Extensive psychometric properties have been documented in the original RF manual (Tellegen, & Ben-Porath, 2008). Van der Heijden and colleagues (2008, 2010) have evaluated the Dutch counterpart scales, and found results that were overall quite similar. Internal consistency estimates (Cronbach's alpha) for this sample ranged from .72 to .87 for the Higher-Order scales, from .64 to .89 for the Restructured Clinical (RC) scales, from .45 to .81 for the Specific Problem scales, and from .61 to .78 for the Personality Psychopathology Five (PSY-5) scales. See Table 2 for Cronbach's alpha's for all scales.

Table 2 *MMPI-2-RF Correlates with the PCL-R Total Score, and Facet Scores (N=127)*

MMPI-2 scales	α	PCL-R Total score	PCL-R Facet 1	PCL-R Facet 2	PCL-R Facet 3	PCL-R Facet 4
Higher Order scales						
EID	.87	.025	-.092	-.076	.104	.043
THD	.72	.062	-.044	.063	.087	.115
BXD	.74	.449***	.167	.190*	.520***	.447***
Restructured Clinical scales						
RCd	.89	-.007	-.088	-.076	.042	.030
RC1	.81	-.035	-.162	-.132	.075	.030
RC2	.64	-.048	-.102	-.089	.011	-.037
RC3	.80	.103	-.072	.101	.119	.144
RC4	.69	.509***	.229**	.209*	.524***	.539***
RC6	.70	.138	.012	.133	.120	.160
RC7	.81	-.076	-.165	-.184*	.003	.007
RC8	.64	.051	-.058	-.009	.097	.094
RC9	.74	.213*	.098	.080	.202*	.165
Specific Problems: Somatic/Cognitive						
MLS	.68	-.046	-.140	-.126	.068	-.012
GIC	.81	.026	-.040	-.111	.117	.084
HPC	.71	-.086	-.154	-.106	-.012	-.061
NUC	.68	.018	-.107	-.024	.102	.050
COG	.73	.027	-.093	-.022	.109	.053

(Table continues)

(Table 2 continued)

MMPI-2 scales	α	PCL-R Total score	PCL-R Facet 1	PCL-R Facet 2	PCL-R Facet 3	PCL-R Facet 4
Specific Problems: Internalizing scales						
SUI	.48	.082	-.049	.064	.119	.107
HLP	.64	.065	-.036	.099	.108	.017
SFD	.70	.020	-.041	-.040	.064	.036
NFC	.70	.059	-.057	.004	.133	.040
STW	.65	.014	-.047	-.082	.039	.069
AXY	.57	-.118	-.165	-.154	-.024	-.042
ANP	.74	.202*	.104	.109	.234**	.126
BRF	.45	-.093	-.141	-.153	-.006	-.002
MSF	.68	-.007	.023	-.084	.024	.024
Specific Problems: Externalizing scales						
JCP	.74	.458***	.187*	.213*	.511***	.495***
SUB	.47	.323***	.084	.071	.387***	.355***
AGG	.66	.202*	.135	.087	.162	.191*
ACT	.60	.075	-.036	.001	.111	.068
Specific Problems: Interpersonal scales						
FML	.76	-.008	-.050	-.048	-.011	.025
IPP	.69	-.302***	-.299***	-.276**	-.229**	-.163
SAV	.72	-.084	-.143	-.052	-.034	-.044
SHY	.76	-.239**	-.318***	-.228**	-.112	-.165
DSF	.66	-.130	-.174	-.096	-.116	-.071
Interest scales						
AES	.56	.091	.194*	.082	.037	-.005
MEC	.46	-.049	.000	.072	-.072	-.032
Personality Psychopathology Five						
AGGR-r	.69	.348***	.333***	.275**	.237**	.230**
PSYC-r	.71	.046	-.084	.020	.095	.096
DISC-r	.61	.378***	.118	.201*	.457***	.359***
NEGE-r	.78	.086	.008	-.066	.114	.097
INTR-r	.72	-.075	-.061	-.033	-.068	-.064

Note: EID = Emotional/Internalizing Dysfunction, THD = Thought Dysfunction, BXD = Behavioral/Externalizing Dysfunction, RCD = Demoralization, RC1 = Somatic Complaints, RC2 = Low Positive Emotions, RC3 = Cynicism, RC4 = Antisocial Behavior, RC6 = Ideas of Persecution, RC7 = Dysfunctional Negative Emotions, RC8 = Aberrant Experiences, RC9 = Hypomanic Activation, MLS = Malaise, GIC = Gastrointestinal Complaints, HPC = Head Pain Complaints, NUC = Neurological Complaints, COG = Cognitive Complaints, SUI = Suicidal/Death Ideation, HLP = Helplessness/Hopelessness, SFD = Self-Doubt, NFC = Inefficacy, STW = Stress/Worry, AXY = Anxiety, ANP = Anger Proneness, BRF = Behavior-Restricting Fears, MSF = Multiple Specific Fears, JCP = Juvenile Conduct Problems, SUB = Substance Abuse, AGG = Aggression, ACT = Activation, FML = Family Problems, IPP = Interpersonal Passivity, SAV = Social Avoidance, SHY = Shyness, DSF = Disaffiliativeness, AES = Aesthetic-Literary Interests, MEC = Mechanical-Physical Interests, AGGR-r = Aggressiveness-Revised, PSYC-r = Psychoticism- Revised, DISC-r = Disconstraint-Revised, NEGE-r = Negative Emotionality/Neuroticism-Revised, INTR-r = Introversion/Low Positive Emotionality-Revised.

*$p \leq .05$. **$p \leq .01$. ***$p \leq .001$

PCL-R

The PCL-R (Hare, 1991, 2003) consists of 20 items, which can be scored 0 (*definitely does not apply*), 1 (*may apply or partly applies*) or 2 (*definitely applies*), leading to a possible maximum score of 40. The PCL-R is usually scored on the basis of a combination of file information and an extensive interview. It is not a risk assessment instrument per se, but research shows psychopathy levels to be strongly associated with past and future antisocial and violent behavior (Hare et al., in press), specifically the Lifestyle and Antisocial Facets (e.g., Leistico et al., 2008). Extensive psychometric properties have been documented in the manual (Hare, 2003).

No complete formal inter-rater reliability estimates are available for the full present sample. However, inter-rater reliability for a partially (26.7%, or 34 cases) overlapping sample (Hildebrand et al., 2002) based on the same pairs of raters has been estimated previously. The single measure ICC was .88 for the PCL-R total score. In general, ICCs were good to excellent at the individual item-level (Mdn = .67, range .46 to .80). Internal consistency estimates were similar to those reported by Hildebrand et al. (2002), with a Cronbach's alpha of .86 for the PCL-R total score, and a mean inter-item correlation of 0.24 (ranging from -.07 to .66). Internal consistency estimates for the separate facets were lower but still acceptable, with Cronbach's alpha's ranging from .68 to .81 and all inter-item correlations positive (mean inter-item correlations ranging from .14 to .60).

SIDP-IV

The SIDP-IV (SIDP-IV; Pfohl et al., 1997; Dutch translation by De Jong et al., 1996) was administered to assess DSM-IV personality disorder (PD) symptoms. The SIDP-IV follows a topically arranged format (work, interpersonal relations, impulse control, etc.) yielding symptom scores on a 0 (*absent*) to 3 (*strong presence*) scale that are combined into the ten DSM-IV dimensional counts of PD symptoms. Its general psychometric properties are well established (Widiger, 2002). Inter-rater reliability was assessed in a Dutch psychiatric (opiate dependent) sample (n = 50) and showed reasonable to good average criterion reliability with kappa coefficients ranging between 0.76 for schizotypal personality disorder and 0.93 for avoidant personality disorder (Damen et al., 2005). Raters in the current study were extensively trained mental health professionals. Internal consistency estimates in this sample were adequate, with Cronbach's alpha's ranging from .55 to .77 and the majority of estimates falling in the range of .65 to .75.

Results

Table 2 reports zero order correlations between MMPI-2-RF substantive (i.e., all scales except validity) raw scale scores and PCL-R total and facet scores.
The PCL-R total score was highly correlated with Higher-Order scale BXD (Behavioral/ Externalizing Dysfunction). For the Restructured Clinical (RC) scales, PCL-R total score was highly correlated with Antisocial Behavior (RC4) and modestly with Hypomanic Activation (RC9). At the level of the Specific Problems scales, PCL-R total score was modestly correlated with Internalizing scale Anger Proneness (ANP), highly correlated with Externalizing scales Juvenile Conduct Problems (JCP) and Substance Abuse (SUB),

and modestly correlated with Externalizing scale Aggression (AGG). PCL-R total score was also negatively correlated with two of the Interpersonal scales: highly with Interpersonal Passivity (IPP) and moderately with Shyness (SHY). With regard to the Personality Psychopathology Five (PSY-5) scales, PCL-R total score was highly correlated with Aggressiveness-revised (AGGR-r) and Disconstraint-revised (DISC-r).

PCL-R Facet 1 (Interpersonal) was not correlated with any of the Higher Order scales, and moderately correlated with RC4. There was a modest correlation with Externalizing Specific Problems scale JCP, and high negative correlations with Interpersonal scales IPP and SHY. With regard to the PSY-5 scales PCL-R Facet 1 was highly correlated with AGGR-r. PCL-R Facet 2 (Affective) was modestly correlated with Higher Order scale BXD. At the level of the RC scales, there was a modest correlation with RC4, as well as a modest negative correlation with Dysfunctional Negative Emotions (RC7). Facet 2 was modestly correlated with Externalizing Specific Problems scale JCP. Furthermore, there were moderate negative correlations with Interpersonal scales IPP and SHY. Finally, Facet 2 was moderately correlated with PSY-5 scale AGGR-r and modestly with DISC-r.

PCL-R Facet 3 (Lifestyle) was highly correlated with Higher Order scale BXD. With regard to the Restructured Clinical scales, there was a high correlation with RC4 and a modest correlation with RC9. At the level of the Specific Problems scales there was a moderate correlation with Internalizing scale ANP, high correlations with Externalizing scales JCP and SUB, and a moderate negative correlation with Interpersonal scale IPP. Finally, Facet 3 was highly correlated with PSY-5 scale DISC-r and moderately with AGGR-r. PCL-R Facet 4 (Antisocial) was highly correlated with Higher Order scale BXD, and with RC4. With regard to the Specific Problems scales, there were high correlations with Externalizing scales JCP and SUB, as well as a modest correlation with AGG. Furthermore, Facet 4 was highly correlated with PSY-5 scale DISC-r and moderately with AGGR-r.

None of the PCL-R scores were correlated with any of the Somatic/Cognitive or Interest Scales, except for a very small correlation between PCL-R Facet 1 and Aesthetic-Literary Interests (AES).

Based on the correlations reported above, exploratory stepwise regression (criteria: entry $p < .05$, removal $p > .10$) analyses were conducted by scale set (Higher-Order scales, Restructured Clinical scales, Specific Problems scales and PSY-5 scales) to determine which of the scales from each of these four scale-sets would provide the most information in relation to predicting PCL-R total scores and facet scores. Tables 3 through 7 present the results of these analyses. Results will be discussed by PCL-R (total and facet) score and MMPI-2-RF scale set.

Table 3 *Regressions for PCL-R Total Score (N=127)*

Scale set/block	MMPI-2 RF scale	Standardized β	R	R^2	R^2_{adj}	R^2_{chg}	F Test F	df	p	F_{chg} Analysis F_{chg}	p
Higher Order											
1	BXD		.449	.202	.195		31.553	1, 125			
Restructured Clinical											
1	RC4		.509	.259	.253		43.720	1, 125			
2	RC4	.569	.554	.307	.296	.048	27.480	2, 124	<.001	8.587	.004
	RC7	-.227									
Specific Problems											
1	JCP		.458	.209	.203		33.119	1, 125			
2	JCP	.453	.512	.263	.251	.053	22.073	2, 124	<.001	8.927	.003
	SHY	-.230									
3	JCP	.373	.566	.320	.304	.058	19.302	3, 123	<.001	10.409	.002
	SHY	-.278									
	SUB	.257									
Personality Psychopathology Five											
1	DISC-r		.378	.143	.136		20.801	1, 125			
2	DISC-r	.307	.456	.208	.195	.066	16.306	2, 124	<.001	10.268	.002
	AGGR-r	.266									

Table 4 *Regressions for PCL-R Facet 1 (N=127)*

Scale set/block	MMPI-2 RF scale	Standardized β	R	R²	R²_adj	R²_chg	F Test F	df	p	F_chg Analysis F_chg	p
Higher Order											
1	-										
Restructured Clinical											
1	RC4		.229	.052	.045		6.896	1, 125			
2	RC4	.293	.328	.107	.093	.055	7.449	2, 124	<.001	7.636	.007
	RC7	-.243									
Specific Problems											
1	SHY		.318	.101	.094		14.080	1, 125			
2	SHY	-.315	.366	.134	.120	.033	9.587	2, 124	<.001	4.679	.032
	JCP	.181									
Personality Psychopathology Five											
1	AGGR-r		.333	.111	.104		15.580	1, 125			

Table 5 *Regressions for PCL-R Facet 2 (N=127)*

Scale set/block	MMPI-2 RF scale	Standardized β	R	R²	R²_adj	R²_chg	F Test F	df	p	F_chg Analysis F_chg	p
Higher Order											
1	BXD		.190	.036	.028		4.688	1, 125			
Restructured Clinical											
1	RC4		.209	.044	.036		5.720	1, 125			
2	RC4	.277	.324	.105	.091	.062	7.294	2, 124	<.001	8.524	.004
	RC7	-.257									
3	RC4	.248	.385	.148	.127	.043	7.126	3, 123	<.001	6.180	.014
	RC7	-.354									
	RC6	.234									
Specific Problems											
1	IPP		.276	.076	.069		10.340	1, 125			
2	IPP	-.248	.324	.105	.091	.029	7.282	2, 124	<.001	3.978	.048
	JCP	.172									
Personality Psychopathology Five											
1	AGGR-r		.275	.076	.068		10.241	1, 125			

Table 6 *Regressions for PCL-R Facet 3 (N=127)*

Scale set/block	MMPI-2 RF scale	Standardized β	R	R²	R²adj	R²chg	F Test F	df	p	Fchg Analysis Fchg	p
Higher Order											
1	BXD		.520	.270	.264		46.269	1, 125			
Restructured Clinical											
1	RC4		.524	.274	.268		47.206	1, 125			
Specific Problems											
1	JCP		.511	.261	.256		44.254	1, 125			
2	JCP	.433	.565	.320	.309	.058	29.125	2, 124	<.001	10.598	<.001
	SUB	.253									
3	JCP	.395	.593	.352	.336	.032	22.239	3, 123	<.001	6.081	.015
	SUB	.278									
	IPP	-.183									
Personality Psychopathology Five											
1	DISC-r		.457	.209	.203		33.009	1, 125			

CHAPTER 5 — USING MMPI-2-RF CORRELATES TO ELUCIDATE THE PCL-R AND ITS FOUR FACETS IN A SAMPLE OF MALE FORENSIC PSYCHIATRIC PATIENTS

79

Table 7 *Regressions for PCL-R Facet 4 (N=127)*

Scale set/block	MMPI-2 RF scale	Standardized β	R	R²	R²adj	R²chg	F Test F	df	p≤	F chg Analysis F chg	p≤
Higher Order											
1	BXD		.447	.199	.193		31.129	1, 125			
Restructured Clinical											
1	RC4		.539	.290	.285		51.114	1, 125			
Specific Problems											
1	JCP		.495	.245	.239		40.531	1, 125			
2	JCP	.426	.539	.290	.279	.045	25.355	2, 124	<.001	7.931	.006
	SUB	.224									
3	JCP	.409	.575	.330	.314	.040	20.224	3, 123	<.001	7.360	.008
	SUB	.265									
	SHY	-.204									
Personality Psychopathology Five											
1	DISC-r		.359	.129	.122		18.485	1, 125			

For the PCL-R total score (Table 3), Higher-Order scale BXD accounted for 20.2% of the variance, RC scales RC4 and (low) RC7 combined to account for 30.7% of the variance, Specific Problems scales JCP, IPP (negatively), and SUB accounted for 32.0% of the variance, and PSY-5 scales, DISC-r and AGGR-r combined to account for 20.8% of the variance. For Facet 1 (Table 4), no higher order scales emerged as significant predictors (although a trend was notable), RC scales RC4 and (low) RC7 combined to account for 10.7% of the variance, Specific Problems scales (low) SHY and JCP combined to account for 13.4% of the variance, and PSY-5 scale AGGR-r accounted for 11.1% of the variance. For Facet 2 (Table 5), Higher-Order scale BXD accounted for 3.6% of the variance, RC scales RC6, (low) RC7, and RC4 combined to account for 14.8% of the variance, Specific Problems scales IPP (negatively) and ANP accounted for 10.5% of the variance, and PSY-5 scale AGGR-r accounted for 7.6% of the variance. For Facet 3 (Table 6), Higher Order scale BXD accounted for 27.0% of the variance, RC scale RC4 accounted for 27.4% of the variance, Specific Problems scales JCP, SUB, and IPP (negatively) accounted for 35.2% of the variance, and PSY-5 scale DISC-r accounted for 20.9% of the variance. Finally, for facet 4 (Table 7), Higher Order scale BXD accounted for 19.9% of the variance, RC scale RC4 accounted for 29.0% of the variance, Specific Problems scales JCP and SUB accounted for 33.0% of the variance, and PSY-5 scale DISC-r combined to account for 12.9% of the variance.

Discussion

The current investigation had two specific aims: a) juxtaposing the full MMPI-2-RF psychopathology and personality models with the Four-Facet PCL-R model, and in so doing b) elucidate the nature of the comprising facets of psychopathy. To our knowledge, this is the first study to test these associations in a group of male forensic psychiatric patients.

Overall, zero order correlations were largely in line with previous research (Anderson et al., 2015; Philips et al.,2013; Sellbom, 2011; Sellbom et al., 2012; Sellbom et al., 2015; Sellbom et al., 2005; Sellbom et al., 2007; Wygant & Sellbom, 2012). Furthermore, our sets of exploratory regression analyses predicting psychopathy as measured by the PCL-R yield a consistent picture at each level of the MMPI-2-RF hierarchy (see Table 8 for a summary of these findings). In male forensic psychiatric patients, Behavioral/Externalizing Dysfunction (BXD) robustly predicted psychopathy and its facets. An interesting pattern of results emerged at the RC level. Antisocial behavior as measured by RC4 was consistently predictive of (global) psychopathy and its comprising facets. Dysfunctional negative emotions (low RC7) was also predictive of PCL-R total score, but only of its Interpersonal and Affective Facets, not the behavioral Facets 3 and 4. In a similar vein, the PSY-5 indices of abnormal personality displayed a conceptually meaningful pattern. Specifically, disinhibition and instrumental aggression (as measured by DISC-r and AGGR-r) were jointly predictive of the PCL-R total score, but only DISC-r predicted the Lifestyle and Antisocial Facets, while AGGR-r predicted the Affective and Interpersonal facets. The Specific Problems scales display a very interesting pattern of correlates and predictors with the PCL-R. No Somatic/Cognitive scales were related to

Table 8 *Summary of Results of Stepwise Regression Analyses Predicting PCL-R Total and Facet Scores*

| PCL-R | MMPI-2-RF | | | |
	H-O scales	RC scales	SP scales	PSY-5
Total score	BXD	RC4, RC7(-)	JCP, SHY (-), SUB	DISC-r, AGGR-r
Facet 1 (interpersonal)	-	RC4, RC7(-)	SHY (-), JCP	AGGR-r
Facet 2 (affective)	BXD	RC4, RC7(-), RC6	IPP (-), JCP	AGGR-r
Facet 3 (impulsive)	BXD	RC4	JCP, SUB, IPP (-)	DISC-r
Facet 4 (antisocial)	BXD	RC4	JCP, SUB, SHY (-)	DISC-r

Note: H-O = Higher Order scales, RC = Restructured Clinical scales, SP = Specific Problems Scales, PSY-5 = Personality Psychopathology Five scales, BXD = Behavioral/Externalizing Dysfunction, RC4 = Antisocial Behavior, RC6 = Ideas of Persecution, RC7 = Dysfunctional Negative Emotions, JCP = Juvenile Conduct Problems, SUB = Substance Abuse, IPP = Interpersonal Passivity, SHY = Shyness, AGGR-r = Aggressiveness-Revised, DISC-r = Disconstraint-Revised.

the PCL-R (or added to the prediction of its scores), and several conceptually relevant associations and predictors were found in Externalizing and Interpersonal domains. Juvenile Conduct Problems (JCP) was predictive of the PCL-R total score and each of its facets, while Substance Abuse (SUB) only added to the prediction of the behavioral facets. Low scores on Interpersonal Passivity (IPP) were associated with the PCL-R total Score and added uniquely to the prediction of its Affective and Lifestyle Facets. Low Shyness (SHY) added to the prediction of PCL-R total Score, as well the Interpersonal and Antisocial Facets.

While the present findings are overall largely consistent with the extant body of evidence, it deserves mention that previous studies found stronger relationships between low scores on the internalizing MMPI-2-RF scales and (components of) psychopathy. More specifically, apart from RC7, SHY and IPP, we did not find any other negative associations with internalizing scales, including NEGE-r, INTR-r, EID, Demoralization (RCd), and several of the Specific Problems scales, as was observed in some of the previous studies. Several factors might account for these differential findings, and we offer some speculations. First, there is great diversity in sample composition, with some studies including students (together with forensic participants), all studies included female participants, and one study included females only. Moreover, these studies differed in terms of their operation-alization of psychopathy, with most studies relying on self-report indices (mostly PPI; Philips et al., 2013; Sellbom, et al, 2005; Sellbom, et al., 2012; Sellbom et al., 2015). Of course, these studies are vulnerable to inflated correlations due to shared method variance. Finally, our all male sample was in mandatory treatment for various forms of externalizing behavior, which sets them apart from some of the other studies.

It is instructive to compare and contrast our results using the PCL-R psychopathy as a dependent measure, with those observed in a different sample from the same Dutch clinic, predicting DSM-defined Antisocial Personality Disorder (ASPD; Anderson et al., 2015). Across studies, a robust association between RC4 and both ASPD and psychopathy was observed. However, whereas ASPD was predicted by antisocial behavior and impulsivity (RC4 and RC9), psychopathy as measured by the PCL-R, was best predicted by antisocial behavior and a lack of negative emotionality (RC4 and RC7), which is much in line with early theorizing by Cleckley (1941/1976). In clinical practice, this may be a useful distinction. For example, clinicians encountering patients with a high score on RC4, may inspect scores on RC7 and RC9 to aid the differentiation between ASPD and psychopathy.

Another notable finding was the specific predictive contribution of RC6 (Ideas of Persecution) to the Affective Facet of the PCL-R (Facet 2). Previous research has demonstrated that Facet 2 is associated with less therapeutic change and more drop-out in forensic treatment settings (Olver et al., 2013; Olver & Wong, 2011), while not a strong predictor of recidivism (Yang et al., 2010). Interestingly, Rock and colleagues (2013) reported a comparable finding using MMPI-2-RF estimates of psychopathy, calculated from the conceptual perspective of the PPI. In their sample of 483 convicted male batterers undergoing treatment, Fearless Dominance was not related to recidivism but did add to the prediction of treatment failure. Apparently, Facet 2 complicates treatment, and the association with RC6 yields a new hypothesis regarding possible mechanisms. Earlier theorizing from Olver and Wong (2011) hypothesized that affective and empathy deficits possibly undermine the development of insight and the willingness to truly engage in treatment. Hence, it may be the callous and unemotional traits that cause difficulty in establishing an effective therapeutic alliance, and the clinician is advised to focus on working collaboratively towards well-defined goals using a cognitive-behavioral approach while avoiding too much focus on the therapeutic bond. The current finding that RC6 is related to Facet 2 suggests that interpersonal alienation, suspiciousness, and the belief that others seek to harm you may also play a role in not engaging in treatment. RC6 is also related to lack of insight and the tendency to blame others, and is associated with interpersonal difficulties (Ben-Porath, 2012). Alienation and suspiciousness are not explicitly assessed by the PCL-R. However, other authors have noted indications that distrust plays a role in psychopathy (Hildebrand & De Ruiter, 2004; Cooke et al., 2012). If indeed interpersonal alienation and suspiciousness are part of the problem in the treatment of psychopathic patients with high scores on Facet 2, this would call for certain adaptations of the treatment program. To minimize the treatment interfering effect of distrust, for example, treatment providers would need to pay special attention to transparency about treatment methods and goals, accountability during the treatment process, and clarity about other rules and expectations.

There are a number of limitations to the current study that warrant consideration. First and foremost, we used a naturalistic sample of all male forensic psychiatric patients. Cross validation of several key findings in different samples is indicated. For example, future research may examine whether RC4, RC7 and RC9 may be useful to differentiate

ASPD and psychopathy in incarcerated offenders, or among general offenders under supervision in the community. Furthermore, this study needs replication in similar high-risk samples, as well as in female offenders. Finally, our exploratory analytic strategy using stepwise regression analyses amplifies the need for cross-validation of the emergent predictors. The general consistency of findings across facets, however, mitigates this concern.

In conclusion, the current investigation demonstrated both discriminant and convergent validity across MMPI-2-RF scale sets and the Four-Facet model of the PCL-R. Additionally, the current study provides suggestive evidence how the MMPI-2-RF (i.e., RC6 elevations) may aid the identification of a specific subgroup of psychopathic offenders that has been shown to be at higher risk for treatment failure and drop-out. If this is replicated, the combination of the MMPI-2-RF and the PCL-R may prove very valuable in identifying these particularly challenging patients, allowing clinicians to adapt treatment more effectively.

References

Anderson, J. L., Sellbom, M., Pymont, C., Smid, W., De Saeger, H., & Kamphuis, J. H. (2015). Measurement of DSM–5 Section II personality disorder constructs using the MMPI–2–RF in clinical and forensic samples. *Psychological Assessment, 27*(3), 786-800. https://doi.org/10.1037/pas0000103

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

Ben-Porath, Y. S. (2012). *Interpreting the MMPI-2-RF.* University of Minnesota Press.

Ben-Porath, Y. S., & Tellegen, A. (2008). *MMPI-2-RF (Minnesota Multiphasic Personality Inventory-2-Restructured Form): Manual for administration and scoring.* University of Minnesota Press.

Benning, S. D., Patrick, C. J., Hicks, B. M., Blonigen, D. M., & Krueger, R. F. (2003). Factor structure of the psychopathic personality inventory: validity and implications for clinical assessment. *Psychological Assessment, 15*(3), 340-350. https://doi.org/10.1037/1040-3590.15.3.340

Cooke, D. J., Hart, S. D., Logan, C., & Michie, C. (2012). Explicating the construct of psychopathy: Development and validation of a conceptual model, the Comprehensive Assessment of Psychopathic Personality (CAPP). *International Journal of Forensic Mental Health, 11*(4), 242-252. https://doi.org/10.1080/14999013.2012.746759

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment, 13*(2), 171-188. https://doi.org/10.1037/1040-3590.13.2.171

Cleckley, H. (1976). *The mask of sanity: An attempt to reinterpret the so-called psychopathic personality.* Mosby. (Original work published 1941)

Damen, K.F.M., De Jong, C. A. J., & Van der Kroft, P. J. A. (2005). Interrater reliability of the Structured Interview for DSM-IV Personality in an opioid-dependent patient sample. *European Addiction Research,10*(3), 99–104. https://doi.org/10.1159/000077697

Derksen, J., De Mey, H., Sloore, H., & Hellenbosch, G. (1993). *MMPI–2: Handleiding bij afname, scoring en interpretatie* [MMPI–2: Manual for administration, scoring, and interpretation]. PEN Test Publishers.

Forsman, M., Lichtenstein, P., Andershed, H., & Larsson, H. (2010). A longitudinal twin study of the direction of effects between psychopathic personality and antisocial behaviour. *Journal of Child Psychology and Psychiatry, 51*(1), 39-47. https://doi.org/10.1111/j.1469-7610.2009.02141.x

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

Hare, R. D. (2003). *The Hare Psychopathy Checklist-Revised: 2nd edition.* 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., & Mokros, A. (in press). The PCL-R assessment of psychopathy: Development, properties, debates, and new directions. In: C. J. Patrick (Ed.), *Handbook of psychopathy* (2nd ed.). The Guilford Press.

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

Hildebrand, M., & de Ruiter, C. (2004). PCL-R psychopathy and its relation to DSM-IV Axis I and II disorders in a sample of male forensic psychiatric patients in the Netherlands. *International Journal of Law and Psychiatry, 27*(3), 233-248. https://doi.org/10.1016/j.ijlp.2004.03.005

Hildebrand, M., de Ruiter, C., de Vogel, V., & van der Wolf, P. (2002). Reliability and factor structure of the Dutch language version of Hare's Psychopathy Checklist-Revised. *International Journal of Forensic Mental Health, 1*(2), 139-154. https://doi.org/10.1080/14999013.2002.10471169

Jong, C.A.J. de, Derks, F.C.H., Oel, C.J. van & Rinne, Th. (1996). *SIDP-IV: Gestructureerd interview voor de DSM-IV persoonlijkheidsstoornissen* [*SIDP-IV: Structured Interview for DSM-IV Personality*]. Stichting Verslavingszorg Oost Brabant.

Karpman, B. (1946). Psychopathy in the scheme of human typology. *The Journal of Nervous and Mental Disease, 103*(3), 276-288. https://psycnet.apa.org/doi/10.1097/00005053-194603000-00007

Leistico A. M., 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

Lilienfeld, S. O., & Andrews, B. P. (1996). Development and preliminary validation of a self-report measure of psychopathic personality traits in noncriminal populations. *Journal of Personality Assessment, 66*(3), 488-524. https://doi.org/10.1207/s15327752jpa6603_3

Miller, J. D., & Lyman, D. R. (2014). Psychopathy and personality: Advances and debates. *Journal of Personality, 83*(6), 585-592. https://doi.org/10.1111/jopy.12145

Mitchell, I. J., Smid, W.J. , Troelstra, J., Wever, E., Ziegler, T. E., & Beech, A. R. (2013). Psychopathic characteristics are related to high basal urinary oxytocin levels in male forensic patients. *Journal of Forensic Psychiatry & Psychology, 24*(3), 309-318. https://doi.org/10.1080/14789949.2013.773455

Neumann, C. S., Hare, R. D., & Pardini, D. A. (2015). Antisociality and the construct of psychopathy: Data from across the globe. *Journal of Personality, 83*(6), 678-692. https://doi.org/10.1111/jopy.12127

Neumann, C. S., & Pardini, D. (2014). Factor structure and construct validity of the Self-Report Psychopathy (SRP) Scale and the Youth Psychopathic Traits Inventory (YPI) in young men. *Journal of Personality Disorders, 28*(3), 419-433. https://doi.org/10.1521/pedi_2012_26_063

Olver, M. E., Lewis, K., & Wong, S. C. (2013). Risk reduction treatment of high-risk psychopathic offenders: The relationship of psychopathy and treatment change to violent recidivism. *Personality Disorders: Theory, Research, and Treatment, 4*(2), 160-167. https://doi.org/10.1037/a0029769

Olver, M. E., & Wong, S. (2011). Predictors of sex offender treatment dropout: Psychopathy, sex offender risk, and responsivity implications. *Psychology, Crime & Law, 17*(5), 457-471. https://doi.org/10.1080/10683160903318876

Patrick, C. J., Edens, J. F., Poythress, N. G., Lilienfeld, S. O., & Benning, S. D. (2006). Construct validity of the psychopathic personality inventory two-factor model with offenders. *Psychological Assessment, 18*(2), 204-208. https://doi.org/10.1037/1040-3590.18.2.204

Pfohl, B., Blum, N., & Zimmerman, M. (1997). *Structured Interview for DSM-IV Personality (SIDP-IV)*. American Psychiatric Press.

Phillips, T. R., Sellbom, M., Ben-Porath, Y. S., & Patrick, C. J. (2013). Further development and construct validation of MMPI-2-RF indices of global psychopathy, fearless-dominance, and impulsive-antisociality in a sample of incarcerated women. *Law and Human Behavior, 38*(1), 34-46. https://doi.org/10.1037/lhb0000040

Polaschek, D. L., & Daly, T. E. (2013). Treatment and psychopathy in forensic settings. *Aggression and Violent Behavior, 18*(5), 592-603. https://doi.org/10.1016/j.avb.2013.06.003

Poythress, N. G., Lilienfeld, S. O., Skeem, J. L., Douglas, K. S., Edens, J. F., Epstein, M., & Patrick, C. J. (2010). Using the PCL-R to help estimate the validity of two self-report measures of psychopathy with offenders. *Assessment, 17*(2), 206-219. https://doi.org/10.1177/1073191109351715

Rock, R.C., Sellbom, M., Ben-Porath, Y.S., & Salekin, R.T. (2013). Concurrent and predictive validity of psychopathy in a batterers intervention sample. *Law and Human Behavior, 37*(3), 145-154. https://doi.org/10.1037/lhb0000006

Sellbom, M. (2011). Elaborating on the construct validity of the Levenson Self-Report Psychopathy Scale in incarcerated and non-incarcerated samples. *Law and Human Behavior, 35*(6), 440-451. https://doi.org/10.1007/s10979-010-9249-x

Sellbom, M., Ben-Porath, Y. S., Lilienfeld, S. O., Patrick, C. J., & Graham, J. R. (2005). Assessing psychopathic personality traits with the MMPI-2. *Journal of Personality Assessment, 85*(3), 334-343. https://doi.org/10.1207/s15327752jpa8503_10

Sellbom, M., Ben-Porath, Y. S., & Stafford, K. S. (2007). A comparison of MMPI-2 measures of psychopathic deviance in a forensic setting. *Psychological Assessment, 19*(4), 430-436. https://doi.org/10.1037/1040-3590.19.4.430

Sellbom, M., Ben-Porath, Y. S., Patrick, C. J., Wygant, D. B., Gartland, D. M., & Stafford, K. P. (2012). Development and construct validation of MMPI-2-RF indices of global psychopathy, fearless-dominance, and impulsive-anti-sociality. *Personality Disorders: Theory, Research, and Treatment, 3*(1), 17-38. https://doi.org/10.1037/a0023888

Sellbom, M., Drislane, L. E., Johnson, A. K., Goodwin, B. E., Phillips, T. R., & Patrick, C. J. (2015). Development and validation of MMPI-2-RF scales for indexing triarchic psychopathy constructs. *Assessment, 23*(5), 527-543. https://doi.org/10.1177/1073191115590853

Skeem, J. L., & Cooke, D. J. (2010). Is criminal behavior a central component of psychopathy? Conceptual directions for resolving the debate. *Psychological Assessment, 22*(2), 433-445. https://doi.org/10.1037/a0008512

Tellegen, A., Ben-Porath, Y. S., McNulty, J. L., Arbisi, P. A., Graham, J. R., & Kaemmer, B. (2003). *MMPI-2 Restructured Clinical (RC) Scales: Development, validation, and interpretation.* University of Minnesota Press.

Tellegen, A., & Ben-Porath, Y. S. (2008). *MMPI-2-RF (Minnesota Multiphasic Personality Inventory-2-Restructured Form): Technical manual.* University of Minnesota Press.

Van der Heijden, P. T., Egger, J. I. M., & Derksen, J. J. L. (2008). Psychometric evaluation of the MMPI–2 Restructured Clinical Scales in two Dutch samples. *Journal of Personality Assessment, 90*(5), 456-464. https://doi.org/10.1080/00223890802248745

Van der Heijden, P. T., Egger, J. I. M., & Derksen, J. J. L. (2010). Comparability of scores on the MMPI–2–RF scales generated with the MMPI–2 and MMPI–2–RF booklets. *Journal of Personality Assessment, 92*(3), 254-259. https://doi.org/10.1080/00223891003670208

Widiger, T. A. (2002). Personality disorders. In: M. M. Antony & D. H. Barlow (Eds.), *Handbook of assessment and treatment planning for psychological disorders* (pp. 453-480). The Guilford Press.

Wygant, D. B., & Sellbom, M. (2012). Viewing psychopathy from the perspective of the Personality Psychopathology Five model: Implications for DSM-5. *Journal of Personality Disorders, 26*(5), 717-726. https://doi.org/10.1521/pedi.2012.26.5.717

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

Do we Hold Males and Females to the Same Standard? A Measurement Invariance Study on the Psychopathy Checklist-Revised

Abstract

Psychopathy in females has been understudied. Extant data on gender comparisons using the predominant measure of assessment in clinical practice, the Psychopathy Checklist Revised (PCL-R), points to a potential lack of measurement invariance (MI). If indeed the instrument does not perform equally (well) in both genders, straightforward comparison of psychopathy scores in males and females is unwarranted. Using a sample of female and male forensic patients (N=110 and N=147 respectively), we formally tested for MI in a structural equation modeling framework. We found that the PCL-R in its current form does not attain full MI. Four items showed threshold-biases and particularly Factor 2 (the Social Deviance Factor) is gender biased. Based on our findings, it seems reasonable to expect that specific scoring adjustments might go a long way in bringing about more equivalent assessment of psychopathic features in men and women. Only then can we begin to meaningfully compare the genders on the prevalence, structure, and external correlates of psychopathy.

Psychopathy is considered a highly relevant syndrome in forensic mental health. Its association with violent behavior and criminal recidivism is well established (Hare & Neumann, 2008; Leistico et al., 2008), and it has been found to be predictive of poor treatment response (Ogloff et al., 1990; Wong & Hare, 2005). The most widely used instrument to assess psychopathy in clinical practice is the Psychopathy Checklist-Revised (PCL-R; Hare, 1991, 2003), and a host of studies documented favorable psychometric properties in a variety of samples of predominantly male violent (sexual) offenders. Psychopathy in females was initially substantially understudied, but in the past two decades this situation has improved (Verona & Vitale, 2018). The extant PCL-R research in females yields a picture of possible gender differences in reliability and particularly in validity. A crucial implication of the notion that the instrument may, psychometrically, not perform equally (well) in both genders, is that straightforward comparison of males and females with psychopathy may in fact be unwarranted.

Several studies provided indicative evidence for differences in the psychometric properties of the PCL-R across genders. First, ten years after the publication of the first edition of the PCL-R (Hare, 1991), Vitale and Newman (2001) published a review of studies on the psychometric performance of the PCL-R in females. The authors concluded that the PCL-R appeared to have good inter-rater reliability and internal consistency when used in female populations. The results regarding the base rate of psychopathy among women varied widely across female samples. Of particular interest was the observation that when restricted to incarcerated women, observed psychopathy prevalences were generally lower than in males, with estimates ranging from 11-23% in women to 15-30% in men. The authors discussed several possible reasons for this finding, one of which was that the PCL-R may not adequately capture psychopathy in women. An alternative explanation was that some of the items of the PCL-R may not have the same sensitivity for females. Moreover, there were indications for gender differences in factor structure. Exploratory factor analysis in male samples originally led to the identification of two underlying factors, a model that was reliably replicated (Hare & Neumann, 2008). Factor 1 contains the personality traits typically associated with psychopathy; i.e., shallow affect, lack of empathy, and a manipulative, arrogant interpersonal style. Factor 2 reflects a chronically impulsive, aggressive and antisocial lifestyle. Preliminary evidence indicated that the factor structure was different in females; that is, not all individual items loaded equally across genders. However, because they used a small sample of mixed ethnicity, Vitale and Newman remained cautious, and recommended the use of factor analysis in larger samples, as well as Item Response Theory (IRT) analysis to compare structure and functioning of items across gender.

Bolt and colleagues took up this challenge and conducted multigroup IRT analyses in several populations, described as "female criminal offenders," "male criminal offenders," and "male forensic psychiatric patients" (Bolt et al., 2004). An assumption in IRT is that there is a single latent factor or trait underlying all items of the respective scale. IRT can detect differences in item-trait functioning across groups. Bolt and colleagues argued that the extant factor models of the PCL-R all have highly intercorrelated factors, and

verified that a single-factor model adequately fit their data. Items of the PCL-R that showed differential item functioning across genders were items 5 (Conning/manipulative), 12 (Early behavior problems), 18 (Juvenile delinquency), and 20 (Criminal versatility), with females scoring higher on item 5, and lower on items 12, 18 and 20, respectively. Bolt and colleagues then proceeded their analyses using the original Two-Factor model, as well as the more fine-grained Four-Facet model that Hare introduced in the second edition of the PCL-R (Hare, 2003) to discuss the results. Factor 1 (consisting of the Interpersonal and Affective Facets 1 and 2) did not display large differences between male and female offenders. Evidence for differential item functioning was primarily found in the behavioral items of Factor 2. Of the facets residing under Factor 2, Facet 4 (Antisocial; consisting of items 10, 12, 18, 19, and 20) in particular produced consistently lower scores for females. Bolt and colleagues also computed the information function for the four facets and the Total Score. The information function indicates the precision with which the facets measure the latent traits. In females, Facets 3 (Lifestyle) and 4 (Antisocial) provided notably less information regarding the latent traits than in males. The authors pointed out that not only were the behaviors measured with these facets likely to be less prevalent among females, they were also more weakly related to psychopathy. Nevertheless, based on the information function of the PCL-R Total Score, their conclusion was that the effect on test-functioning was quite modest, especially at the high-end scores. Accordingly, they deemed the PCL-R (Total Score) effective for the task of distinguishing psychopathic from non-psychopathic females.

A more recent systematic review (Beryl et al., 2014) again evaluated the current evidence regarding prevalence and factor structure of psychopathy in female populations in secure settings, defined as both criminal justice and secure inpatient settings. Of 28 studies that were included, 21 used the PCL-R. Heterogeneity in sample characteristics and study procedures prevented quantitative data synthesis, but prevalence in females was again found to be generally lower than in male populations in secure settings. Nine of the reviewed studies investigated factor structure of the PCL-R. Although two studies replicated the Two-Factor/ Four-Facet model that was found in males, Beryl and colleagues concluded that Cooke and Michie's Three-Factor model (2001) had best fit for females. In this model, items related to criminal and antisocial behavior were removed; the remaining three factors were named 'Arrogant/Deceitful Interpersonal Style', 'Deficient Affective Experience', and 'Impulsive/Irresponsible Behavioural Style' (Cooke & Michie, 2001). These three factors are identical to the first three facets of the Four-Facet model, and consist of 13 of the 20 original items of the PCL-R. Cooke and Michie argued that antisocial behavior is a correlate, and possibly better thought of as a possible consequence of psychopathy, rather than a core feature. The debate about whether or not antisocial behavior should be part of the conceptualization of psychopathy is ongoing. However, considering the IRT-findings by Bolt and colleagues (2004) that illuminated that Facet 4 appears to function less well for females, it is perhaps not surprising that the Three-Factor model yielded best fit.

To our knowledge, there are no other studies comparing the structure of the PCL-R across gender in adult samples. However, two studies utilizing the Psychopathy Checklist: Youth Version (PCL:YV; Forth et al., 2003) may also be informative (Dillard et al., 2013; Tsang et al., 2015). The PCL:YV was modeled on the PCL-R, and although several items were adapted to the specific population of adolescents, it appears to have a similar Two-Factor/Four-Facet structure (Hare, et al., 2018). Dillard and colleagues (2013) used IRT to compare functioning of the PCL:YV in adolescent boys (n = 307) and girls (n = 144) that had come into contact with the law. For girls, quite a few items (11 of 20: items 1, 3, 4, 6, 7, 11, 12, 13, 16, 18, and 20) demonstrated differential functioning when compared to boys. The information function of all four facets was higher for boys than for girls. However, comparable to Bolt et al. (2004), Factor 1 (incorporating Facet 1 and 2) was less sensitive to gender than Factor 2 (Facet 3 and 4). In general, Factor 1 was more useful in identifying the underlying construct of psychopathy than Factor 2. Tsang et al. (2015) used IRT with Cooke and Michie's Three-Factor model in a larger sample of adolescents involved in the justice system (N = 1007, 38% female). Of the 13 items, 4 showed differential functioning across gender (2, 13, 14, and 16), partly overlapping with the results of Dillard et al. (2013). In general, Cooke and Michie's Factor 1 and 2 (identical to Facet 1 and 2) were again found to be better at discriminating high levels of the underlying trait of psychopathy, than Factor 3.

Taken together, the observed gender differences in prevalence, factor structure, and item functioning, albeit rather inconsistent across studies, point to the possibility that the PCL-R does not possess the same psychometric properties across gender. This is problematic, as this would imply that the scores on the PCL-R cannot be meaningfully compared across males and females. To be able to meaningfully compare male and female scores, the PCL-R should adhere to measurement invariance across gender (Mellenbergh, 1989; Meredith, 1993). Measurement invariance implies that a male and a female with the same position on a factor underlying the PCL-R (e.g., 'lack of empathy') should have the same expected score on the item(s) measuring that factor (see for example Eigenhuis et al., 2017). If measurement invariance is violated, males and females that are (for instance) equally empathic will nevertheless display differences in their empathy scores. As a result, differences in the item scores between males and females do not necessarily indicate that there are 'real' differences in the underlying pathology. Therefore, establishing measurement invariance is important to guarantee the comparability of male and female scores of the PCL-R. Additionally, it is a prerequisite to be able to compare predictive validity across gender.

In conclusion, there appears to be only one study testing measurement invariance of the PCL-R across gender in an adult sample (Bolt et al., 2004). Of course, replication is needed in different samples and settings to be able to come to more robust conclusions. As Bolt et al. (2004) only considered female criminal offenders, male criminal offenders, and male forensic psychiatric patients, the aim of the present study is to test for measurement invariance with respect to gender in a forensic psychiatric population of males and females. Key objective is to clarify possible gender biases in the functioning of the PCL-R

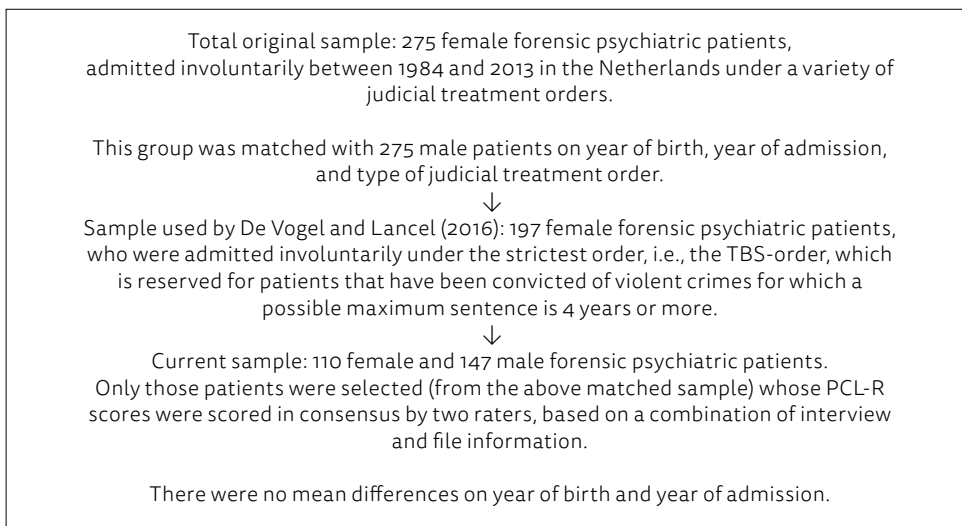in forensic psychiatric samples, with the overall aim to inform and improve current clinical practice.

Method

Data

In 2012, a multicenter research project was started in the Netherlands on gender differences in forensic psychiatry (for more information, see De Vogel et al., 2016). The project and all studies connected to it (including the current study) were conducted with official permission from the directors of the hospitals involved, in compliance with the (then) applicable ethical standards. The patients were recruited from four different forensic psychiatric hospitals in The Netherlands. The total female sample (*N* = 275) represented nearly all female forensic psychiatric patients in the Netherlands admitted involuntarily between 1984 and 2013 with a variety of judicial treatment orders. The female sample was matched with a male sample based on year of birth, year of admittance, and type of judicial treatment order.

De Vogel and Lancel (2016) used a subgroup of 197 female and 197 male patients to study gender differences in the assessment and manifestation of psychopathy. All patients were admitted involuntarily under the same treatment order, the so-called TBS-order (*ter beschikking stelling*). This type of TBS-order is reserved for patients who have been convicted of violent crimes for which a possible maximum sentence is 4 years in prison. For a more elaborate description of the TBS-system, see Klein Haneveld et al. (2020). The data from the respective hospitals were aggregated, as allocation to the different hospitals had been random and there were no mean differences between the patients

Figure 1 *Flowchart of Sample Composition*

Total original sample: 275 female forensic psychiatric patients, admitted involuntarily between 1984 and 2013 in the Netherlands under a variety of judicial treatment orders.

This group was matched with 275 male patients on year of birth, year of admission, and type of judicial treatment order.
↓
Sample used by De Vogel and Lancel (2016): 197 female forensic psychiatric patients, who were admitted involuntarily under the strictest order, i.e., the TBS-order, which is reserved for patients that have been convicted of violent crimes for which a possible maximum sentence is 4 years or more.
↓
Current sample: 110 female and 147 male forensic psychiatric patients. Only those patients were selected (from the above matched sample) whose PCL-R scores were scored in consensus by two raters, based on a combination of interview and file information.

There were no mean differences on year of birth and year of admission.

from the different hospitals on nationality and IQ. The full range of possible PCL-R scores was included. However, the sample was not matched on administration and scoring of the PCL-R. For 87 females and 50 males the rating was not done in consensus, and partly based on file information only. See De Vogel and Lancel (2016) for more details about the data. For the current study, see our flowchart of sample composition (Figure 1). We used a sub-group of the matched sample described above, with 110 female and 147 male patients. Only those patients were selected whose PCL-R scores were scored in consensus by two raters, based on a combination of interview and file information. In this sub-group, there were no mean differences on year of birth and year of admission.

Instrument
The PCL-R (Hare, 1991, 2003) consists of 20 items, which can be scored 0 (*definitely does not apply*), 1 (*may apply or partly applies*), or 2 (*definitely applies*), leading to a maximum Total Score of 40. The rating is based on a review of extensive file information, preferably in combination with a semi-structured interview. Psychometric properties have been documented in the current manual (Hare, 2003). In previous research at one of the hospitals involved in this study, significant inter-rater reliability was established (single measure ICC = .88 for Total Score) for a partially overlapping sample (Hildebrand et al., 2002).

*Analytic Strategy*

Factor Structure of Psychopathy
In this study, we rely on confirmatory factor analysis of the PCL-R which is, like the IRT approach by Bolt et al. (2004), a suitable approach to test for measurement invariance across gender (Putnick & Bornstein, 2016). To this end, a theory about the underlying PCL-R factor structure is required. At this point we note that in addition to a factor structure, psychopathy can also be represented as a network of symptoms (e.g., Verschuere et al., 2018). We consider a network perspective valuable as it can be used to clarify what the central features of psychopathy are. However, here we focus on quantifying individual differences using the PCL-R. Therefore, similar to, for example, Bolt et al. (2004) and Hare and Neumann (2008), we consider factor analysis a suitable approach.

Since the introduction of the PCL-R several factor structures have been proposed. In the one-factor model (Bolt et al., 2004), referred to as "1F", the only source of individual differences is assumed to be a General Psychopathy Factor. On the contrary, the Two-Factor model (Hare, 1991), referred to as "2F", identifies an Interpersonal/Affective Factor (Factor 1) and a Social Deviance Factor (Factor 2). In addition, in the Four-Facet model (e.g., Hare, 2003), referred to as "4F", a distinction is made between an Interpersonal, an Affective, a Lifestyle, and an Antisocial Facet. A hierarchical model has been proposed which combines the 4F with the 2F at the higher level. That is, the higher-order model contains four-first order facets and two second-order factors (Hare & Neumann, 2005) and is referred to as "4F-2S". In this model, the interpretation of the first- and second-order factors is the same as in the 4F and the 2F above. The idea of the General Psychopathy Factor in the 1F has also been added to the 4F, which results in a

model with four first-order facets and a single second-order factor, referred to as "4F-1S" (Neumann, Hare, & Newman, 2007). Again, the interpretation of these factors is the same as in the 4F and 1F.

The PCL-R contains two items (items 11 and 17) that do not explicitly measure one of the lower-order factors in the models above. However, to enable analyses of these two items we follow the idea presented by Bolt et al. (2004) and include the two items in an overarching higher-order factor. This factor thus includes all lower-order factors and items 11 and 17 as indicators. For model 1F, this simply means that the two items are included as indicators of the single (and only) factor. For the 2F model, this implies that a higher-order factor is added that includes the 2 factors and items 11 and 17 as indicators. The resulting model thus coincides with the 2F-1S model. For the 4F-2S model, we added a third-order factor (making the model a 4F-2S-1T model) that includes the 2 second-order factors and items 11 and 17 as indicators. Thus, the models that we considered are the 1F, the 2F-1S, the 4F-1S, and the 4F-2S-1T model. For the 4F-2S-1T, a graphical representation of this model is depicted in Figure 2.

Figure 2 *Graphical Representation of the 4F-2S-1T Model*
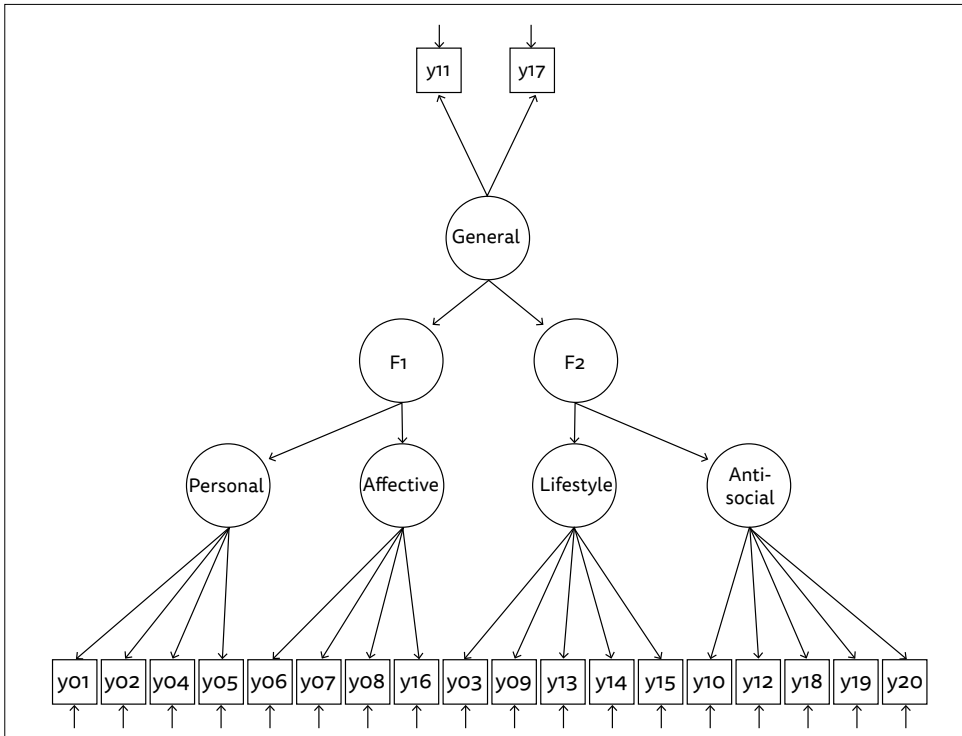
Note. 'Personal' denotes the Interpersonal Factor, 'F1' denotes the Interpersonal/Affective Factor, 'F2' denotes the Social Deviance Factor, and 'General' denotes the General Psychopathy Factor.

Note that we have chosen to use the full PCL-R, and have not included Cooke and Michie's Three-Factor model as a separate model. Testing the MI of the PCL-R according to the Cooke and Michie model would mean leaving out 7 of the 20 items. The seven items that fall outside the model are considered very relevant for clinical forensic practice, as they mostly address different forms of antisocial behavior. Furthermore, the three Cooke and Michie factors were incorporated in identical form into the more fine-grained Four-Facet model of the PCL-R. The present results using the four facets are informative about the three factors of Cooke and Michie. However, at the suggestion of a reviewer, we did verify our results in the hierarchical Three-Factor model; please see footnotes 3 and 4.

Identification of the Models

To enable identification, we included a number of constraints. Specifically, in the 2F-1S model, the two second-order loadings (of the two first-order factors on the second-order factor) were constrained to be equal. In the 4F-2S-1T model, the first-order factors that load on the same second-order loadings were constrained to be equal (i.e., the second-order loadings of the first-order Facets 1 and 2 were constrained to be equal, and the second-order loadings of the first-order Facets 3 and 4 were constrained to be equal). In addition, the third-order loadings of second-order factor 1 and 2 on the third-order factor were also constrained to be equal. All models were identified by fixing the factor means to 0 and the factor variances to 1 within the male and female sample. For the model that fits the data best, we tested for measurement invariance across gender.

General Procedure to Test for Measurement Invariance Across Gender

For measurement invariance to be tenable, the factor model parameters discussed above should be the same in the male and female sample while allowing for group differences in the mean factor scores and the variance of the factor scores (Meredith, 1993; Mellenbergh, 1989). As discussed in the introduction section, most of the previous studies have compared the PCL-R across males and females in terms of psychometric properties like reliability, predictive validity, or classification rates. However, these psychometric properties do not account for possible differences in factor means and variances, therefore, differences in these properties are hard to interpret in terms of violations of measurement invariance.

To assess measurement invariance, we fitted explicit confirmatory factor models to the PCL-R data of the males and females in which factor model parameters are explicitly separated from the factor means and variances.
Specifically, we considered the following increasingly restrictive models: For, *configural invariance* (Horn & McArdle, 1992), the factor loadings and item thresholds are free to vary across groups.[1] If this model shows acceptable fit to the data, it indicates that the same factor structure holds in both groups. For *metric invariance* (sometimes referred to as weak factorial invariance; Widaman & Reise, 1997), the factor loadings are constrained to be equal across groups whereas the factor variances are allowed to differ across groups (allowing free estimation of item thresholds across groups). Should the metric invariance model fit better than the configural model, this would indicate that the constraints

imposed in the model (equal factor loadings across groups) are statistically tenable. In the third, *full measurement invariance* model, the item thresholds are constrained across groups and the factor means are allowed to freely differ across groups (Meredith, 1993). Note that in this final model, the factor variances are still allowed to differ across groups. When the full measurement invariance fits better than the metric model, it in turn reflects that the constraints imposed (equal item thresholds across groups) are statistically tenable.

The procedure outlined above can be repeated at the second-order level to establish invariance of the first-order factors (see Chen et al., 2005). In this procedure, the first-order factor means take over the role of the item thresholds above. That is, as a first step (second-order metric invariance), the second-order loadings are constrained and the second-order factor variances are freed across groups. In the next step (second-order full measurement invariance), the first-order factor means are constrained to be equal across groups while allowing for a mean difference on the second-order factor. Similarly, this procedure can be extended to test for measurement invariance at the level of the third-order factor. Although you can also test for higher-order variances to be equal across groups in principle (Chen et al., 2005) we followed Dolan et al. (2006) and left the higher-order residuals to be free across groups.

Model Fit Assessment and Diagnosis of Misfit
To assess model fit, various fit indices are available. Here, we used the Root Mean Squared Error of Approximation (RMSEA; Browne, & Cudeck, 1989), the Comparative Fit Index (CFI; Bentler, 1990), and the Tucker-Lewis Index (Tucker & Lewis, 1973). For the RMSEA, it holds that values between 0.08 and 0.05 indicate acceptable model fit, and values smaller than 0.05 indicate good model fit. For CFI and TLI these cut-off values are 0.95-0.97 for acceptable model fit, and larger than 0.97 for good model fit (see Schermelleh-Engel et al., 2003).

In the process of fitting these series of models to data, it may be that a given restriction (e.g., the restriction of equal factor loadings) results in a deterioration in model fit indicating that at least one factor loading is not equal across groups. Such sources of misfit can be diagnosed using so-called "modification indices" which are available for each constrained parameter. In the case of a large modification index (a commonly used cut-off value is 10 as this is the default in Mplus; Muthén & Muthén, 1998-2017), the corresponding parameter (in this case the factor loading) can be freed across groups.

---

1 Note that we treat the item scores explicitly as ordinal (which is equivalent to using the graded response model as Bolt et al., 2004, did.). As a result, similarly as in the graded response model, the factor models considered here do not include intercept parameters, but category thresholds (as there are three answer categories, each item contains two thresholds). In addition, the models considered do not contain residual variances. However, given appropriate identification restrictions (see Millsap & Yun-Tein, 2004) both the intercepts and the residual variances can be estimated and tested for invariance. Here however, we fixed the residual variances and intercepts to be equal across groups in all models, and we only freed specific residual variance parameters or intercepts parameters if this is indicated by the modeling results (i.e., poor model fit and/or large modification indices for the residual variances). Note that if the thresholds and factor loadings are shown to be invariant, it can be concluded that the intercepts and residual variances are also invariant.

If the resulting model fits better than the previous model (i.e., the model with all factor loadings to be free across groups), the measurement invariance tests may be continued. If for a minority of the items measurement invariance is violated, (i.e., a few items have factor loadings or item thresholds that are non-invariant across groups) this is referred to as partial measurement invariance (Byrne et al., 1989). The latent variable means and variances under partial measurement invariance can be meaningfully interpreted, but the non-invariant item thresholds and factor loadings do not contribute to the means and variances respectively.

All analyses were conducted in Mplus (Muthén & Muthén, 2017). The observed item scores were explicitly treated as ordinal. Parameters were estimated using weighted least squares estimation. The resulting models are therefore not linear factor models but discrete factor models (Takane & De Leeuw, 1987; Wirth & Edwards, 2007).

Results

Factor Models for the PCL-R
First, we fit the four competing PCL-R models to the data of the males and females. For the 4F-2S-1T, we fixed the variance of first-order Lifestyle Facet (Facet 3) in the female sample to 0 as this variance parameter approached 0 during parameter estimation. This implies, for this model, that the second-order Social Deviance Factor (F2) explains all the variance in the first-order Lifestyle Facet. Results concerning the fit of the four models is depicted in Table 1. As can be seen, all models but the one factor model showed acceptable fit according to the RMSEA statistic. However, according to the CFI and TLI values all models fit poorly. Of note, the CFI and TLI are incremental fit indices, meaning that they measure the added value of the corresponding model to a baseline model in which all variables are uncorrelated. As generally correlations are medium to small in non-cognitive data (e.g., our average absolute inter-item correlation is 0.240), not much incremental fit can be expected. We therefore focused on the RMSEA, and inspected the CFI and TLI values only to compare subsequent models. As can be seen from the table, for the models considered, the Four-Facet model with two second-order factors and

Table 1 *Fit Indices for the Four PCL-R Models Considered in this Study*

Model	par	χ^2(df)	RMSEA	CFI	TLI
1F	120	641.221 (340)	0.083 (0.073; 0.093)	0.842	0.823
2F-1S	122	522.771 (338)	0.065 (0.054; 0.076)	0.903	0.891
4F-1S	128	517.760 (332)	0.066 (0.055; 0.077)	0.902	0.888
4F-2S-1T[a]	126	484.981 (334)	0.059 (0.047; 0.071)	0.921	0.910

[a] Variance of Facet 3 fixed to 0 in the female sample

one third-order factor (4F-2S-1T) is associated with the best RMSEA (0.059). Additional inspection of the results for this model indicated that there are no obvious sources of misfit (the largest modification index was 11.582 for the residual covariance between items 1 and 2). Therefore, we accepted the 4F-2S-1T model and tested for measurement invariance with respect to gender in this model.

Measurement Invariance Analyses

The results concerning measurement invariance are in Table 2. As discussed above, we started with the configural invariance model (Model 1), which is the same as the 4F-2S-1T model in Table 1. Model fit was considered acceptable as judged by the RMSEA. Next, we fixed the first-order factor loadings to be equal across groups, while freeing the first-order facet variances in the female group. The resulting model (Model 2) yielded better fit than the configural invariance model in terms of RMSEA, CFI and TLI, indicating that the factor loadings are equal across gender. Next, we constrained the item thresholds to be equal across groups, while freeing the first-order facet means in the female sample. The resulting model (Model 3) fit worse as compared to the previous model as judged by the RMSEA, CFI, and TLI. The modification indices of the item threshold parameters indicated that the item threshold parameters of items 2 and 10 were a significant source of misfit. We freed these parameters resulting in Model 3′. The fit of this model is approximately the same as Model 2 indicating that except for the items 2 and 10 item thresholds were the same across genders. The standardized estimates of the thresholds of these two non-invariant items are depicted in Table 3. As can be seen, the thresholds for item 2 were smaller in the male group while for item 10, the thresholds were smaller in the female group.

Table 2 *Fit Indices for the Models to Establish Measurement Invariance with Respect to Gender*

	Model	par	χ^2(df)	RMSEA	CFI	TLI
1	Configural invariance[a]	126	484.980 (334)	0.059 (0.047; 0.071)	0.921	0.910
2	1st-order loadings	111	484.593 (349)	0.055 (0.043; 0.066)	0.929	0.922
3	item thresholds	79	549.531 (381)	0.059 (0.047; 0.069)	0.911	0.912
3′	item thresholds[b]	83	520.061 (377)	0.054 (0.042; 0.065)	0.925	0.924
4	2nd-order loadings	83	521.733 (377)	0.055 (0.043; 0.066)	0.924	0.923
5	1st-order means	81	556.910 (379)	0.060 (0.049; 0.071)	0.907	0.906
5′	1st-order means[c]	82	523.819 (378)	0.055 (0.043; 0.066)	0.923	0.923
6	3rd-order loadings	80	537.171 (380)	0.057 (0.045; 0.068)	0.917	0.917
6′	3rd-order loadings[d]	81	521.137 (379)	0.054 (0.042; 0.065)	0.925	0.925
7	2nd-order means	77	559.161 (383)	0.060 (0.049; 0.070)	0.907	0.908
7′	2nd-order means[e]	79	526.830 (381)	0.055 (0.043; 0.066)	0.923	0.924

[a]Variance of Facet 3 fixed to 0 in the female sample. [b]Item thresholds of item 2, and item 10 free across groups. [c]Mean of first-order Facets 3 and 4 free in the female sample, mean of second-order factor 2 fixed to 0 in the female sample. [d]Factor loading of item 11 free across groups. [e]Item thresholds of item 17 free across groups.

Table 3 *Standardized Parameter Estimates for the Non-Invariant Thresholds in Model 3′*

Item	Threshold	Estimate	SE	Z	P
		Males			
2	1	-0.546	0.109	-5.002	0.000
	2	0.691	0.113	6.120	0.000
10	1	-1.181	0.136	-8.669	0.000
	2	-0.150	0.105	-1.421	0.155
		Females			
2	1	0.321	0.143	2.242	0.025
	2	1.268	0.202	6.265	0.000
10	1	-1.812	0.192	-9.436	0.000
	2	-0.692	0.175	-3.962	0.000

As partial measurement invariance was now established with respect to the first-order facet structure, the first-order facet means and variances can now be meaningfully compared across groups. See Table 4 for the standardized parameters estimates of the first-order facet means and variance in the female sample. As the means in the males are all equal to 0 for identification purposes, the standardized estimates for the first-order facet means in the table can readily be interpreted as standardized mean differences between males and females. See Figure 3 for a graphical display of these differences. As can be seen, female patients scored significantly lower than their male counterparts on all facets, with the largest standardized effect on the Antisocial Facet (Facet 4) and the smallest effect on the Interpersonal Facet (Facet 1). In addition, the female sample showed less variance on all first-order facets.

Table 4 *Standardized Parameter Estimates of the First-Order Facet Means and Variances in the Female Sample in Model 3′*

Factor	Estimate	SE	Z	P
	Means			
1: Interpersonal Facet	-0.549	0.174	-3.163	0.002
2: Affective Facet	-0.689	0.164	-4.194	0.000
3: Lifestyle Facet	-0.612	0.175	-3.489	0.000
4: Antisocial Facet	-1.498	0.255	-5.885	0.000
	Variances			
1: Interpersonal Facet	0.212	0.147	1.443	0.149
2: Affective Facet	0.457	0.111	4.097	0.000
3: Lifestyle Facet	0.000	-	-	-
4: Antisocial Facet	0.370	0.154	2.412	0.016

Note. Means in the male sample are fixed to 0 for all first-order facets. In addition, variances in the male sample are fixed to 1.00 for all first-order facets. The variance of first-order Facet 3 is fixed to 0 in the female sample as this variance approached 0 in model estimation.

Figure 3 *Graphical Display of the Standardized Mean Differences in the First-Order Facets in Model 3'*
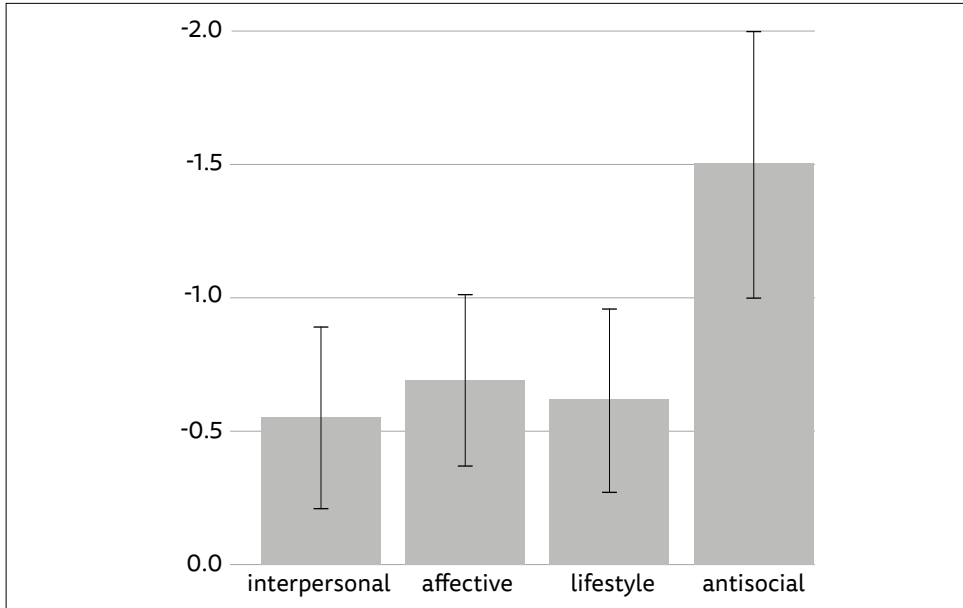
Note. Vertical lines represent 95% confidence intervals.

Invariance of the Second-Order Structure

Next, in Model 3' we forced the second-order factor loadings to be equal across gender while freeing the variances of the two second-order factors (Model 4). As can be seen in Table 1 the fit of this model is identical to the previous model (Model 3') as the model contains the same number of parameters to model the correlations among the first-order factors.[2] We can thus accept Model 4 and fix the first-order means to be equal to 0 (i.e., equal across groups) while allowing for a mean difference on the second-order factors. The resulting model (Model 5) fits worse as compared to Model 4 as judged by the RMSEA, CFI, and TLI. The modification indices of the first-order factor mean parameters indicated that the means of the Lifestyle Facet (Facet 3) and the Antisocial Facet (Facet 4) cannot be explained by the mean difference in the second-order Social Deviance Factor only. We therefore freed the first-order facet means of the Lifestyle Facet and the Antisocial Facet and fixed the mean of second-order Social Deviance Factor to be equal to 0. The resulting model (Model 5') showed largely the same fit to the data as Model 4 with the RMSEA, and TLI being identical for Model 4 and 5', and with the CFI being 0.001 smaller for Model 5'. As this difference in CFI is negligible, and no misfit was evident from the modification indices (all modification indices are smaller than 10), we accepted Model 5'. In this model, the first-order facet means are thus modeled using one second-order mean parameter (standardized estimate: -0.816, SE: 0.193), and mean parameters for the first-order Lifestyle Facet (standardized estimate: -0.593, SE: 0.170) and the first-order Antisocial

---

2 As the second-order factor loadings of first-order Facets 1 and 2 and the second-order factor loadings of Facets 3 and 4 are constrained to be the same, see above, we only had to fix 2 parameters, i.e., the two loadings, and we had to free 2 parameters, i.e., the second-order factor variances

Facet (standardized estimate: -1.467, SE: 0.243). As indicated by the estimates of the first-order means, the females scores on these facets are smaller than that of the males.

Invariance of the Third-Order Factor Structure

In Model 5' we proceeded by fixing the third-order factor loading (note that we only have one third-order factor loading as discussed above) and the factor-loadings of item 11 and 17 while allowing for a variance difference on the third-order general psychopathology factor. Note that item 11 and 17 were now also taken into account, as these items are direct indicators of the third-order factor. As can be seen in Table 2, the resulting model (Model 6) fit worse than the previous Model 5' in terms of all fit indices. The modification indices indicated that the misfit was mainly due to the factor loading of item 11. Freeing this parameter across groups resulted in a model (Model 6') that showed better than Model 5' for all fit indices. Result indicated that the factor loading of item 11 is smaller in the male group (standardized estimate: 0.405, se: 0.101) as compared to the female group (standardized estimate 0.735, se: 0.087). Finally, we equated the second-order means and the thresholds of item 11 and 17 to be equal across groups (Model 7). Model fit again deteriorated as compared to the previous Model 6'. The modification indices indicated that the misfit was mainly due to the thresholds of item 17. Freeing these parameters across groups resulted in Model 7' which has the same fit as Model 6'. However, as Model 7' contains fewer parameters, we accepted this model as the final model. In this final model, the thresholds for item 17 in the male group are: 0.859 (se: 0.127) and 1.189 (se: 0.145), while in the female group they are -0.105 (se: 0.157) and 0.202 (se: 0.168). That is, the thresholds were smaller in the female group, indicating that females scored higher on item 17 than would be expected on basis of their underlying third-order general psychopathology factor score. In the final Model 7', the standardized mean difference between males and females on the third-order General Psychopathy Factor equals -0.895 (SE: 0.198). Note that this mean difference is only based on the second-order Interpersonal/Affective Factor (F1) and on item 11, as the second-order Social Deviance Factor (F2) and item 17 were shown to have non-invariant means/thresholds.

Discussion

This study used confirmatory factor analysis within a structural equation modeling framework to test for measurement invariance of the PCL-R for gender, using 147 male and 110 female forensic psychiatric patients. The PCL-R was found to be only partially invariant for gender. More specifically, four items (i.e., items 2, 10, 11, and 17) violated measurement invariance.[3] That is, for item 2 ('Grandiose sense of self-worth') the thresholds were larger in the female group than in the male group, indicating that females scored lower on these items than would be expected based on their position on the underlying first-order Interpersonal Facet. In addition, for items 10 ('Poor behavioral

3   As suggested by a reviewer, we verified these results in the Cooke and Michie (2001) second-order Three-Factor model. Similar to our third-order Four-Facet model, we found item 2 to violate threshold invariance in the Cooke and Michie model. The other items that we found to be non-invariant in the third-order Four-Facet model, items 10, 11, and 17, are not part of the Three-Factor model. The results are available upon request.

controls') and 17 ('Many short-term marital relationships'), the thresholds were smaller in the female group indicating that females scored higher on these items than would be expected based on their position on the underlying facet or factor (which is the first-order Antisocial Facet for item 10, and the General Psychopathy Factor for item 17). With respect to item 11 ('Promiscuous sexual behavior'), we found that the factor loading of this item on the General Psychopathy Factor was larger in the female group as compared to the male group. This indicates that for females, item 11 was a stronger indicator of the General Psychopathy Factor. Moreover, the first-order Lifestyle and Antisocial Facets (Facets 3 and 4) were biased indicators of the second-order Social Deviance Factor (Factor 2).[4] That is, the female scores on the Lifestyle and Antisocial Facets were smaller than what would be expected on the basis of their position on the underlying Social Deviance Factor. Thus, a female patient with a given position on Factor 2 will have a lower expected position on the Lifestyle and Antisocial Facets as compared to a male patient with exactly the same position on Factor 2. Hence, male and female forensic psychiatric patients cannot be meaningfully compared on the Social Deviance Factor, as they are not held to the same standard. For the first-order facet structure partial invariance was established, and accordingly male and female forensic psychiatric patients could be meaningfully compared on the first-order facets after removing items 2 and 10.

It is informative to compare our results to those of Bolt and colleagues (2004) in a group of incarcerated females. At item level our findings do not overlap, as Bolt et al. found differential item functioning in items 5, 12, 18, and 20, whereas we found items 2, 10, 11, and 17 to violate measurement invariance. Possible explanations for these differences are variations in sample (a criminal justice population versus forensic psychiatric patients) and the use of unmatched versus matched comparison groups. However, at the level of the Four-Facet model our findings are remarkably similar, with Facets 3 and 4 (Lifestyle and Antisocial) not functioning well. This may be a possible explanation for the finding that Factor 2, comprising of Facets 3 and 4 and predictive of violence in men (Yang et al., 2010), has not been consistently found to predict violence in women. Possibly, prevalence and predictive validity with regard to violent behavior will improve when the non-invariant items are removed. This would be an interesting follow-up study. Of Factor 1, in both studies only one item violated measurement invariance; item 5 ('Conning/manipulative' in Bolt et al., 2004) and item 2 (Grandiose sense of self-worth in our study). Both are part of Interpersonal Facet 1. Facet 2, relating to defective emotional functioning, often considered one of the core aspects of psychopathy, was found to be invariant. Previously, some authors have speculated whether the construct of psychopathy in females is inherently distinct from psychopathy in males (Miller et al., 2011; Viljoen et al., 2015; Wynn et al., 2012). A tentative conclusion on the basis of these two studies would be that the personality features of psychopathy (the interpersonal

---

4   In the Cooke and Michie (2001) second-order Three-Factor model (see previous footnote), Factor 3 was found to be unbiased with respect to the second-order factor. However, note that the second-order General Psychopathy Factor in the Cooke and Michie model is statistically and substantively different from our second-order Social Deviance Factor in the third-order Four-Facet model.

and affective traits of Facets 1 and 2) are manifested in a sufficiently comparable way in males and females, while the behavioral features (impulsive, irresponsible lifestyle and antisocial behavior) are subject to gender differences. We feel that, although the PCL-R was (again) found not to be fully invariant with respect to gender, it does appear to capture the core features of psychopathy in women.

To speculate on the origin of the partial lack of MI, it may stand to reason that female forensic patients with narcissistic and callous features express their impulsive, irresponsible lifestyle and antisocial behavior systematically different than men with the same trait levels. For instance, several authors have speculated that, due to biological differences, the use of physical force or violence to achieve a desired outcome may be considerably less feasible and effective for women, whereas the manipulative use of flirtation, intimacy or sexual favors may work well to attain their personal agenda (Forouzan & Cooke, 2005; Nicholls & Petrila, 2005; Wynn et al., 2012). The literature which focuses specifically on gender differences in psychopathic behavior still appears to be rather sparse. An early review of behavioral gender differences (among other aspects of psychopathy) by Forouzan and Cooke (2005) reported that impulsivity and conduct disorder in females were characterized by self-harming behavior, manipulation, and complicity in theft and fraud, as opposed to violent behavior in males. This was recently replicated by De Vogel and Lancel (2016). More research is needed that explores typical behaviors associated with psychopathy in women.

If we accept the finding that the PCL-R does not optimally capture female behavioral manifestations, what are we to do with the assessment of females with psychopathic features? One possibility would be to uphold the items of Facets 1 and 2 (Factor 1), and to formulate a new female version of Facets 3 and 4 (Factor 2), based on female manifestations of impulsive, irresponsible, parasitic and antisocial behavior. For example, we found Item 11 (Promiscuous Sexual Behavior) to be more strongly related to the General Psychopathy Factor for females than for males. Item 11 describes sexual behavior that is impersonal and trivial, as reflected in frequent casual contacts, infidelities, and prostitution. In male populations Item 11 does not load on any facets or factors. However, for females this may well be a good candidate for the behavioral features typically associated with Factor 2. A less drastic alternative to rewriting all Factor 2 items would be to develop supplementary guidelines to the existing manual. In a series of studies, Morrissey and colleagues investigated the applicability of the PCL-R in individuals with intellectual disabilities in secure settings (Morrissey, 2003; Morrissey et al., 2005; Morrissey, et al., 2007a; Morrissey et al. 2007b). The authors developed such supplementary guidelines for all PCL-R items, while maintaining the flavor or intent of the original items (Morrissey, 2007). Where appropriate, they expanded item descriptions to include examples of behavior especially relevant for individuals with intellectual disabilities. For some items, the criteria for evidence of the behavior involved were slightly broadened, or different sources of information were added. In a validation study in three forensic settings (N = 203), use of the supplementary guidelines in combination with the PCL-R showed adequate internal consistency (alpha = .82) and good interrater reliability

(ICC = .89) (Morrissey et al., 2005). What would this idea mean for the items we found to be non-invariant for females? First, female forensic psychiatric patients scored lower on item 2 (Grandiose Sense of Self-Worth) than would be expected based on their position on the underlying Interpersonal Facet. Possibly, the current phrasing of the item does not adequately capture female narcissistic manifestations. Perhaps, providing additional examples of grandiosity in females, possibly less overt and brazen, more subtle and covert, may improve the scoring of this item. Second, female psychiatric patients scored higher than would be expected on items 10 (Poor Behavioral Controls). Apparently, we tend to judge females more harshly when it comes to aggressive behavior, while we are more tolerant of such behavior in men. More elaborate specification of what type of behavior is needed for a score of 1 or 2 in females may improve veridical assessment. Finally, females also scored higher than would be expected on item 17 (Many Short-term Marital Relationships). The item description in this case precisely defines the number of live-in relationships needed for a score of 1 or 2, also based on age (under 30, or 30 and above). These norms could quite easily be adjusted for females.

There are a number of notable strengths and limitations to this study. First and foremost, as far as we are aware, our study drew on the largest published database of matched male and female forensic psychiatric patients. The principal limitations are inherent to the matched control design: although the case matching was conducted with great care, it is impossible to preclude hidden confounding variables. This limitation is especially relevant to the current study, as we have used a subset of the matched sample based on (homogeneity in) administration and scoring. Also, in the original matched sample, it was not always possible to match a female case with a male case from the same hospital. Thus, the majority of male cases came from the Van der Hoevenkliniek. It is not immediately clear how this may have affected the findings. Second, our study participants consisted of forensic psychiatric patients, all part of the Dutch TBS-measure. Therefore, the results are drawn from and hence applicable to a very specific forensic population, not necessarily generalizable to other populations. Also, at item level, findings were disparate from those reported by Bolt et al. (2004). Replication and testing the generalizability to other types of samples is warranted. Thirdly, based on the current study, the practical consequences of the partial MI remain unclear. Female forensic patients may, for example, have been unduly burdened with, or escaped consequences of, partly unreliable psychopathy assessments. Another consequence of partial MI may be the inconsistency in predictive validity of the PCL-R in males and females. In several studies that have examined the predictive validity of the PCL-R for re-offending and/or institutional violence in women (De Vogel et al., 2019; De Vogel & Lancel, 2016; Geraghty & Woodhams, 2015; Loucks & Zamble, 2000; Richards et al., 2003; Salekin et al., 1998; Weizmann-Henelius et al., 2015), non-significant to modest relationships were found. Future studies will have to investigate these issues.

A final point of discussion is whether the sample size in our study (110 female and 147 male patients) is sufficient for measurement invariance analyses. There are no sample size recommendations for testing measurement invariance available that can be directly

applied to our study. For instance, Meade and Bauer (2007) pointed out that a sample size of 100 participants per group is sufficient to detect violations of metric invariance, but their study focused on a Three-Factor model with standardized factor loadings around 0.8 and 20 continuous items. Thus, we did meet the requirement of 100 participants per group, but we used a more complex model and our standardized factor loadings were between 0.4 and 0.9. As we did find violations of measurement invariance in the present study, it can be concluded that sample size is at least not too small, but it cannot be ruled out that we missed some small effects on items with small factor loadings. However, these effects are of less practical significance, and the effects that we did find can be trusted.

In sum, to provide a straightforward answer to the question we posed in the title of our article (i.e., Do we hold males and females with psychopathy to the same standard?), we have to conclude that the PCL-R in its current form does not fully attain this essential outcome. On the other hand, it seems reasonable to expect that specific scoring adjustments (especially with respect to the Social Deviance Factor) might go a long way in bringing about more equivalent assessment of psychopathic features in men and women. Only then can we meaningfully compare the genders on the prevalence, structure, and external correlates of psychopathy.

References

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Beryl, R., Chou, S., & Völlm, B. (2014). A systematic review of psychopathy in women within secure settings. *Personality and Individual Differences*, *71*, 185-195. https://doi.org/10.1016/j.paid.2014.07.033

Bolt, D. M., Hare, R. D., Vitale, J. E., & Newman, J. P. (2004). A multigroup item response theory analysis of the Psychopathy Checklist – Revised. *Psychological Assessment*, *16*(2), 155–168. https://doi.org/10.1037/1040-3590.16.2.155

Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *British Journal of Mathematical and Statistical Psychology*, *37*(4), 62–83. https://doi.org/10.1207/s15327906mbr2404_4

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*(3), 456-466. https://doi.org/10.1037/0033-2909.105.3.456

Chen, F. F., Sousa, K. H., & West, S. G. (2005). Teacher's corner: Testing measurement invariance of second-order factor models. *Structural Equation Modeling*, *12*(3), 471-492. https://doi.org/10.1207/s15328007sem1203_7

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment*, *13*(2), 171-188. https://doi.org/10.1037/1040-3590.13.2.171

De Vogel, V., Bruggeman, M., & Lancel, M. (2019). Gender-sensitive violence risk assessment: Predictive validity of six tools in female forensic psychiatric patients. *Criminal Justice and Behavior*, *46*(4), 528-549. https://doi.org/10.1177/0093854818824135

De Vogel, V., & Lancel, M. (2016). Gender differences in the assessment and manifestation of psychopathy: Results from a multicenter study in forensic psychiatric patients. *International Journal of Forensic Mental Health*, *15*(1), 97-110. https://doi.org/10.1080/14999013.2016.1138173

De Vogel, V., Stam, J., Bouman, Y. H., Ter Horst, P., & Lancel, M. (2016). Violent women: A multicentre study into gender differences in forensic psychiatric patients. *The Journal of Forensic Psychiatry & Psychology*, *27*(2), 145-168. https://doi.org/10.1080/14789949.2015.1102312

Dillard, C. L., Salekin, R. T., Barker, E. D., & Grimes, R. D. (2013). Psychopathy in adolescent offenders: An item response theory study of the antisocial process screening device–self report and the Psychopathy Checklist: Youth Version. *Personality Disorders: Theory, Research, and Treatment*, *4*(2), 101-120. https://doi.org/10.1037/a0028439

Dolan, C. V., Colom, R., Abad, F. J., Wicherts, J. M., Hessen, D. J., & Van De Sluis, S. (2006). Multi group covariance and mean structure modeling of the relationship between the WAIS-III common factors and sex and educational attainment in Spain. *Intelligence*, *34*(2), 193-210. https://doi.org/10.1016/j.intell.2005.09.003

Eigenhuis, A., Kamphuis, J. H., & Noordhof, A. (2017). Personality in general and clinical samples: Measurement invariance of the Multidimensional Personality Questionnaire. *Psychological Assessment, 29*(9), 1111-1119. https://doi.org/10.1037/pas0000408

Forouzan, E., & Cooke, D. J. (2005). Figuring out la femme fatale: Conceptual and assessment issues concerning psychopathy in females. *Behavioral Sciences & the Law, 23*(6), 765-778. https://doi.org/10.1002/bsl.669

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

Geraghty, K. A., & Woodhams, J. (2015). The predictive validity of risk assessment tools for female offenders: A systematic review. *Aggression and Violent Behavior, 21*(2), 25-38. https://doi.org/10.1016/j.avb.2015.01.002

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

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

Hare, R. D., & Neumann, C. S. (2005). Structural models of psychopathy. *Current Psychiatry Reports, 7*(1), 57-64. https://doi.org/10.1007/s11920-005-0026-3

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., & Mokros, A. (2018). The PCL-R assessment of psychopathy: Development, properties, debates, and new directions. In C. J. Patrick (Ed.), *Handbook of psychopathy* (2nd ed., pp. 509-528). The Guilford Press.

Hildebrand, M., de Ruiter, C., de Vogel, V., & van der Wolf, P. (2002). Reliability and factor structure of the Dutch language version of Hare's Psychopathy Checklist-Revised. *International Journal of Forensic Mental Health, 1*(2), 139-154. https://doi.org/10.1080/14999013.2002.10471169

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*(3), 117 – 144. https://doi.org/10.1080/03610739208253916

Klein Haneveld, E., Smid, W., Timmer, K., & Kamphuis, J. H. (2021). Clinical appraisals of individual differences in treatment responsivity among patients with psychopathy: A Consensual Qualitative Research study. *Criminal Justice and Behavior.* Advance online publication. https://doi.org/10.1177/0093854820970597

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

Loucks, A. D., & Zamble, E. (2000). *Predictors of criminal behaviour and prison misconduct in serious female offenders.* Retrieved November 15, 2020, from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.559.5450&rep=rep1&type=pdf

Meade, A. W., & Bauer, D. J. (2007). Power and precision in confirmatory factor analytic tests of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(4), 611-635. https://doi.org/10.1080/10705510701575461

Mellenbergh, G. J. (1989). Item bias and item response theory. *International Journal of Educational Research*, *13*(2), 127-143. https://doi.org/10.1016/0883-0355(89)90002-5

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Ps ychometrika*, *58*(4), 525-543. https://doi.org/10.1007/BF02294825

Miller, J. D., Watts, A., & Jones, S. E. (2011). Does psychopathy manifest divergent relations with components of its nomological network depending on gender? *Personality and Individual Differences*, *50*(5), 564-569. https://doi.org/10.1016/j.paid.2010.11.028

Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-cat-egorical measures. *Multivariate Behavioral Research*, *39*(3), 479-515. https://doi.org/10.1207/S15327906MBR3903_4

Morrissey, C. (2003). The use of the PCL-R in forensic populations with learning disability. *British Journal of Forensic Practice*, *5*(1), 20-24. https://doi.org/10.1108/14636646200300003

Morrissey, C. (2007). *Guidelines for assessing psychopathy in offenders with intellectual disabilities using the PCL-R and PCL:SV.* Retrieved November 15, 2020, from https://www.researchgate.net/publication/257890698_Guidelines_for_use_of_the_PCL-R_and_SV_in_adults_with_ID

Morrissey, C., Hogue, T., Mooney, P., Lindsay, W. R., Steptoe, L., Taylor, J., et al., (2005). Applicability, reliability, and validity of the Psychopathy Checklist-Revised in offenders with intellectual disabilities: Some initial findings. *International Journal of Forensic Mental Health*, *4*(2), 207-220. https://doi.org/10.1080/14999013.2005.10471225

Morrissey, C., Hogue, T., Mooney, P., Allen, C. & Johnston, S., Hollin, C., et al. (2007a). The predictive validity of the PCL-R in offenders with intellectual disability in a high secure hospital setting: Institutional aggression. *Journal of Forensic Psychiatry and Psychology*, *18*(1), 1-15. https://doi.org/10.1080/08990220601116345

Morrissey, C., Mooney, P., Hogue, T., Lindsay, W. R., & Taylor, J. (2007b). Predictive validity of psychopathy in offenders with intellectual disability in high security: Treatment progress. *Journal of Intellectual and Developmental Disability, 32*(2), 125-134. https://doi.org/10.1080/13668250701383116

Muthén, L.K., & Muthén, B.O. (1998-2017). *Mplus User's Guide* (8th ed.). Muthén & Muthén

Nicholls, T. L., & Petrila, J. (2005). Gender and psychopathy: An overview of important issues and introduction to the special issue. *Behavioral Sciences & the Law*, *23*(6), 729-741. https://doi.org/10.1002/bsl.677

Neumann, C. S., Hare, R. D., & Newman, J. P. (2007). The super-ordinate nature of the Psychopathy Checklist-Revised. *Journal of Personality Disorders*, *21*(2), 102-117. https://doi.org/10.1521/pedi.2007.21.2.102

Ogloff, J. R., Wong, S., & Greenwood, A. (1990). Treating criminal psychopaths in a therapeutic community program. *Behavioral Sciences & the Law*, *8*(2), 181-190. https://doi.org/10.1002/bsl.2370080210

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review, 41*, 71-90. https://doi.org/10.1016/j.dr.2016.06.004

Richards, H. J., Casey, J. O., & Lucente, S. W. (2003). Psychopathy and treatment response in incarcerated female substance abusers. *Criminal Justice and Behavior, 30*(2), 251–276. https://doi.org/10.1177%2F0093854802251010

Salekin, R. T., Rogers, R., Ustad, K. L., & Sewell, K. W. (1998). Psychopathy and recidivism among female inmates. *Law and Human Behavior, 22*(1), 109–128. https://doi.org/10.1023/A:1025780806538

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*(2), 23-74.

Takane, Y., & De Leeuw, J. (1987). On the relationship between item response theory and factor analysis of discretized variables. *Psychometrika, 52*(3), 393-408. https://doi.org/10.1007/BF02294363

Tsang, S., Schmidt, K. M., Vincent, G. M., Salekin, R. T., Moretti, M. M., & Odgers, C. L. (2015). Assessing psychopathy among justice involved adolescents with the PCL: YV: An item response theory examination across gender. *Personality Disorders: Theory, Research, and Treatment, 6*(1), 22. https://doi.org/10.1037/per0000094

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*(1), 1-10. https://doi.org/10.1007/BF02291170

Verschuere, B., van Ghesel Grothe, S., Waldorp, L., Watts, A. L., Lilienfeld, S. O., Edens, J. F., ... & Noordhof, A. (2018). What features of psychopathy might be central? A network analysis of the Psychopathy Checklist-Revised (PCL-R) in three large samples. *Journal of Abnormal Psychology, 127*(1), 51-65. https://doi.org/10.1037/abn0000315

Verona, E., & Vitale, J. (2018). Psychopathy in women: Assessment, manifestations, and etiology. In C. J. Patrick (Ed.), *Handbook of psychopathy* (2nd ed., pp. 509-528). The Guilford Press.

Viljoen, S., Cook, A. N., Lim, Y. L., Layden, B. K., Bousfield, N. K., & Hart, S. D. (2015). Are psychopathic and borderline personality disorder distinct, or differently gendered expressions of the same disorder? An exploration using concept maps. *International Journal of Forensic Mental Health, 14*(4), 267-279. https://doi.org/10.1080/14999013.2015.1114535

Vitale, J. E., & Newman, J. P. (2001). Using the Psychopathy Checklist-Revised with female samples: Reliability, validity, and implications for clinical utility. *Clinical Psychology: Science and Practice, 8*(1), 117-132. https://doi.org/10.1093/clipsy.8.1.117

Weizmann-Henelius, G., Virkkunen, M., Gammelgård, M., Eronen, M., & Putkonen, H. (2015). The PCL-R and violent recidivism in a prospective follow-up of a nationwide sample of female offenders. *The Journal of Forensic Psychiatry & Psychology, 26*(5), 667-685. https://doi.org/10.1080/14789949.2015.1049192

Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281–324). American Psychological Association. https://psycnet.apa.org/doi/10.1037/10222-009

Wirth, R. J., & Edwards, M. C. (2007). Item factor analysis: current approaches and future directions. *Psychological Methods*, *12*(1), 58. https://doi.org/10.1037/1082-989X.12.1.58

Wong, S. C. P., & Hare, R. D. (2005). *Guidelines for a psychopathy treatment program*. Multi-Health Systems.

Wynn, R., Høiseth, M. H., & Pettersen, G. (2012). Psychopathy in women: Theoretical and clinical perspectives. *International Journal of Women's Health*, *4*, 257-263. http://dx.doi.org/10.2147/IJWH.S25518

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

Discussion

The central question of this thesis was: Can individual differences in psychopathic offenders be used systematically to inform and improve their forensic psychiatric treatment with the goal of reducing recidivism? I will first summarize key findings related to this central theme per study, next present pertinent methodological considerations, to finally turn to the implications for treatment and future research.

Main Findings

All samples in this thesis consisted of patients with a TBS treatment order of indefinite length. TBS (*ter beschikking stelling*; at the disposal of the state) is a measure of mandatory intensive treatment, that can be ordered by the Dutch courts together with a sentence for violent or sexual offences. The primary goal of TBS treatment is to minimize the risk of recidivism while gradually working towards rehabilitation. After finishing a prison sentence, offenders are transferred to one of 11 high-security hospitals in the Netherlands to start treatment. Dropping out is technically impossible, as only the court has the competence to terminate the treatment order. However, treatment can fail in one hospital, leading to transfer to another hospital and a new treatment attempt. Such transfers of treatment served as our operationalization of drop-out.

To pursue my main research question, we employed several different designs to examine individual differences in psychopathy and their impact on treatability. Chapter 2 sets the stage for the study of individual differences in psychopathy by presenting two famous cases, Dick and Perry, the offenders from the non-fiction/true crime book *In Cold Blood* by Truman Capote (1966). Using such material has several advantages: no privacy issues are involved, and anyone who wishes to acquaint herself with the information we used to score the Psychopathy Checklist-Revised (PCL-R; Hare 1991) can simply read the book. On the other hand, it is important to acknowledge that this method is not entirely in accordance with the standard scoring procedure for the PCL-R, which is based on a clinical interview in combination with file information. Although the information compiled and presented by Truman Capote could be seen, at least partially, as file information, a clinical interview was absent. Our clinical impression beforehand was that the two protagonists possibly were psychopathic perpetrators, however, with very different backgrounds and personalities. Coincidentally, in our assessment the cases accrued an identical total score of 31. As expected, the underlying facet profiles differed, especially the Interpersonal and Antisocial Facets (Facet 1 and 4), with Dick scoring higher on the Interpersonal Facet, and Perry on the Antisocial Facet. Of note, these results mirrored the findings of the two studies that investigated subtypes using latent profile analysis (Mokros et al., 2015) and cluster analysis (Olver et al., 2015), respectively in samples of male offenders scoring high on the PCL-R. Both studies found two primary subtypes, a primarily manipulative and an aggressive variant. Mokros and colleagues commented that these two subtypes seem to differ mainly in their preferred method of achieving their goals: one preferring the use of deception and manipulation, the other more through violent and intimidating behavior. Dick and Perry seemed to demonstrate exactly these different preferred strategies in their criminal behavior.

Chapter 3 involved an empirical approach to identify subtypes of psychopathic offenders. We performed latent profile analysis (LPA) in a sample of 190 male violent and/or sexually violent offenders with a TBS treatment order admitted to the same hospital (the Van der Hoevenkliniek). In this study, the sample contained not only high scoring individuals, but covered the full range of PCL-R scores instead. The objective was to investigate whether/ which profiles would appear, to examine whether the high-scoring manipulative and aggressive variants would be among them, and to test whether the emerging profiles would be clinically useful. Three subtypes with psychopathic traits emerged: two moderately psychopathic profiles (called 'sociopathic' and callous-conning' respectively) and only one high-scoring 'prototypical' one. We did not replicate the manipulative and aggressive variants described in Chapter 2. When testing the clinical utility of the three profiles, the most salient finding was the high drop-out rate in the prototypical group (46.5%), which was significantly higher than in the other groups (drop-out rates between 14.3% and 21.1%). For the prototypical group, drop-out occurred most often in the early stages of treatment, before unsupervised leave was granted. Prototypical psychopathic patients who remained in treatment however, appeared to proceed through the treatment phases in much the same way as their less severe sociopathic counterparts, and showed comparable recidivism rates (i.e., showed similar benefit from treatment). Unfortunately, the profiles did not predict any of the treatment variables when we controlled for PCL-R total score. This indicates that the typology was significantly related to increasing PCL-R total scores and the treatment outcomes may have been simply related to the severity of the psychopathic traits in general, rather than to a certain constellation of traits.

The study described above demonstrated that preventing attrition appears to be a first hurdle to overcome in the search for effective treatment for psychopathic offenders. Therefore, Chapter 4 focused on drop-out of the prototypical psychopathic group, with the aim of exploring what individual characteristics distinguish completers from non-completers. Because so little was known about the characteristics of psychopathic patients that drop out of treatment, we opted for a maximally open-ended, qualitative research approach. We interviewed 11 expert clinicians from two TBS hospitals and analyzed our data according to the Consensual Qualitative Research guidelines (CQR; Hill, 2012; Hill et al., 1997, 2005). We asked the clinicians to draw on their experience in working with psychopathic offenders and to select and compare two specific cases with a PCL-R score of 28 or higher, one considered a treatment failure, the other a treatment success. The interview not only covered patient characteristics, but also treatment (provider) characteristics and other factors the participants deemed associated with drop-out/ successful completion. Overall, with regard to patient characteristics, extremely high scores on PCL-R Facets, especially deceitful interpersonal style (Facet 1) and defective affective experience (Facet 2) were thought to impede treatment retention, particularly by their negative impact on motivation and the therapeutic relationship. Of note, for all facets, the informants reported that it was the *degree* to which a certain facet was present that counted. That is, some degree of deception, arrogance, callousness, impulsivity, and aggression were all expected behaviors of these patients during treatment, but only very

high levels were thought to be seriously treatment interfering. Severity again seemed to play a role (as it did in the previous study). Additionally, several factors not related to psychopathy were mentioned by the informants: Patients with a relatively prosocial, supportive network were deemed to fare better, as did patients with comorbid borderline traits, and/or older patients (presumably because of physical deterioration and/or having "calmed down" over the course of life).

In Chapter 5, I switched from a person-centered approach to a variable-centered approach. We investigated associations between the PCL-R Four-Facet model and the MMPI-2-RF scales, using a sample of 127 male TBS patients. The MMPI-2-RF is a (much) more encompassing model of personality and psychopathology that shows remarkable fit with the dominant full spectrum hierarchical taxonomy of psychopathology (HiToP; Hopwood et al., 2020). The MMPI-2-RF is composed of a hierarchical set of scales, including three Higher-Order scales, underlying nine (Restructured) Clinical scales. At the lowest level of the hierarchy, it comprises of 25 Specific Problems and Interest scales. Finally, the Personality Psychopathology Five scales (PSY-5) systematically tap abnormal personality variation, much in line with the trait aspect of the DSM-5 alternative model for personality disorders (AMPD; American Psychiatric Association, 2013). My goal was to use the MMPI-2-RF to elucidate the nature of the four facets. As this had not yet been done, we used an exploratory analytic strategy to predict PLC-R scores from the MMPI-2-RF. Most of the scales that were found to be predictive of the PCL-R were quite straightforward, as in fully in line with conceptually expected associations. For example, at the level of the Restructured Clinical scales, all facets were predicted by the scale Antisocial Behavior (RC4), Instrumental Aggression (PSY-5; AGGR-r) predicted the Interpersonal and Affective Facets (Facet 1 and 2), and Disinhibition (PSY-5; DISC-r) and Substance Abuse (SUB) were predictive of the Lifestyle and Antisocial Facets (Facet 3 and 4). From the viewpoint of individual differences, the most interesting finding was that the Affective Facet (Facet 2) alone was predicted by Ideas of Persecution (RC6). This scale refers to beliefs that others seek to harm you and is associated with suspicion, alienation, and blaming others for your difficulties, characteristics not measured by the PCL-R. It may be worthwhile to develop specialized treatment strategies to address this issue.

Chapter 6 pursued yet another aspect of individual differences: Do we hold male and female psychopathic patients to the same standard? It is by no means self-evident that males and females can be equally well measured and represented by the same instrument, in our case the PCL-R. Hence, we examined the preliminary question of measurement invariance (MI) with respect to psychopathy in females and males when using the PCL-R. Better fidelity in assessment also is a prerequisite for studying treatment responsivity and possible individual differences in females. From a matched sample of 197 female and 197 male TBS patients, we used a subgroup (N=110 and N=147 respectively) whose PCL-R scores were determined in consensus using standard procedure (interview plus file information). Our results indicated that the PCL-R in its current form does not attain full MI. Four items (2, 10, 11, 17) showed threshold-biases and particularly Factor 2 was gender biased. Apparently, females with narcissistic and callous features express

their impulsive, irresponsible lifestyle and antisocial behavior in a different way than men with the same trait levels. As others have noted before, biological differences could (for example) play a role, as the use of physical force or violence to achieve a desired outcome may be considerably less feasible or effective for women, whereas the manipulative use of flirtation, intimacy or sexual favors may work well (Forouzan & Cooke, 2005; Nicholls & Petrila, 2005; Wynn et al., 2012). To improve assessment of psychopathy in females, in our view it is most feasible to develop supplementary guidelines to the existing manual, comparable to the work of Morrissey and colleagues (Morrissey, 2003; Morrissey et al., 2005; Morrissey, et al., 2007a; Morrissey et al. 2007b), who developed such guidelines for individuals with intellectual disabilities. For example, female patients scored lower on Grandiose Sense of Self-worth (Item 2) than would be expected based on their position on the underlying Interpersonal Facet. Possibly, the current phrasing of the item does not adequately capture female narcissistic manifestations. Providing additional examples of grandiosity in females, possibly less overt and brazen, more subtle and covert, may improve the scoring of this item. In a very recent publication Cunliffe, Gacono, and Smith (2021) in fact presented experimental adaptations of several PCL-R items (items 1, 2, 5, 6, 7, 8, 9, and 11). It seems reasonable to expect that such scoring adjustments might go a long way in bringing about more equivalent assessment of psychopathic features in men and women.

Methodological Considerations

Important limitations across the comprising studies are the nature of the sample composition and the limited sample size. Sample composition affects generalizability. Except our case examples in Chapter 2, all our samples consisted of patients with a TBS treatment order, of which all (Chapters 3 and 5) or a majority (Chapters 4 and 6) were recruited from a single hospital. Although all TBS hospitals in the Netherlands are expected to offer the same type of treatment (aimed at reducing recidivism) for the same type of patients, little is known about actual differences in population and treatment culture. Additionally, TBS patients in general are severely disordered offenders with multiple diagnoses, whom we have investigated solely through the 'lens' of psychopathy. Other disorders may have affected our findings, but it is not immediately evident how. Replication is therefore needed, not just in other TBS hospitals, but also in other forensic populations (for example, antisocial offenders participating in treatment programs in prison). Sample size also affected our (quantitive) findings, particularly in our second study (Chapter 3). The sample was large enough to be able to conduct latent profile analysis. However, the subsequent stratification into subgroups ruled out several meaningful group comparisons due to low power and/or violated the assumptions of survival analysis. On the other hand, our final study (Chapter 6), investigating measurement invariance of the PCL-R with respect to gender, drew on the largest published database of matched male and female forensic psychiatric patients that we know of.

Another limitation concerns the lack of a forensic treatment manual or method. My central objective was to investigate whether individual differences can be used to

inform treatment aimed at reducing recidivism. Ideally, to answer such a question, one would prefer to start with a well-described treatment method and subsequently test whether individual differences affect the efficacy of that method. However, as explained in the Introduction, no such manual exists. In our treatment studies (Chapters 3 and 4), treatment methods evolved during the observation period. Also, the study in Chapter 4 focused solely on drop-out. Although limiting drop-out is an important goal in developing effective treatment for these patients, this study did not focus on the effects of treatment on reducing recidivism. Hence, our work must not be mistaken for a formal evaluation of treatment efficacy. Our studies can best be seen as naturalistic, ecologically valid evaluations of possible differences in responsiveness to treatment-as-usual, and our findings may instead be used to inform the development of a specialized treatment method based on current practice in forensic psychiatry.

Implications for Treatment of Male Psychopathic Offenders

As others have concluded before (e.g., Wong & Hare, 2005; Wong & Olver, 2015), any evidence-based forensic treatment program for psychopathy should be based on the Risk, Need, and Responsivity principles (RNR; Andrews & Bonta, 2017), previously discussed in the Introduction. The RNR-model offers a framework for treatment of offenders in general, not specifically for psychopathic patients. However, there is a substantial body of research supporting these principles (see for a review McGuire, 2015), and there are no compelling reasons to argue that a framework that is effective in reducing recidivism for the total group of offenders will not be applicable to the subgroup of psychopathic ones. To briefly re-iterate: the Risk principle implies that forensic treatment is most effective when the dose of treatment is matched to the level of risk, with higher risk implying more intensive treatment; the Need principle states that forensic treatment should primarily target dynamic (i.e., changeable) risk factors, also called criminogenic needs, associated with reoffending (as opposed to psychiatric disorders as classified with the DSM-5); and finally the Responsivity principle refers to the necessity to adapt treatment to the learning style, handicaps, and talents of the offender.

There is ample research in diverse forensic populations elaborating on the Risk and Need principles. An important result of research associated with the Risk principle is the development of instruments for structured assessment of recidivism risk. A wide range of instruments is in use today, from more general instruments focusing on risk for aggression, to more specialized instruments for, for example, sexual offending, stalking, child abuse, and domestic violence. With respect to the Need principle, according to the RNR-model, the most predictive risk factors/criminogenic needs are the so-called 'Central Eight' (Andrews & Bonta, 2017): 1) history of antisocial behavior; 2) antisocial personality pattern (i.e., impulsivity, low anger management, poor problem-solving skills); 3) antisocial cognition (i.e., "the world is a jungle," "she asked for it."); 4) antisocial associates; 5) family/marital circumstances (i.e., conflict, neglect); 6) school/work (lack of involvement/performance); 7) leisure/recreation (lack of involvement/satisfaction); and 8) substance abuse. The Responsivity principle has in general been a relatively understudied area of study. Responsivity can refer to static factors such as intelligence, cultural

background, or autism, but also to dynamic issues such as motivation, a factor that may be influenced by an intervention such as Motivational Interviewing.

In clinical practice, in my experience, psychopathy is often treated as simply an important risk factor (i.e., 'patients with psychopathy recidivate significantly more often than patients without'), and sometimes (also) as a static responsivity issue. Scientific treatment studies often treat psychopathy as a dichotomous factor; as if it is either present or absent (i.e., 'patients with psychopathy drop out significantly more often than patients without'). Accordingly, research may reinforce clinical thinking about psychopathy as a one-dimensional factor, or even a taxon. In contrast, the present body of research explicitly focused on individual differences among psychopathic patients, inspired by personalized medicine (Ng and Weisz, 2016). More specifically, it was aimed at identifying key individual characteristics predictive of treatment response, in the hope that this would contribute to the development of a specialized treatment manual. I now turn to the conclusions and recommendations that, in my view, ensued from this series of studies.

Are Subtypes Based on the PCL-R Useful for Personalized Treatment?
In the first two studies, we used the PCL-R to identify specific profiles, or constellations of psychopathic features that may be predictive of treatment effect. Our case studies illustrated two types already found in two previous, quantitative studies using samples of exclusively high scoring offenders (on the PCL-R); one manipulative and one aggressive subtype. However, we did not replicate these two high-scoring subtypes in our second study, in which we used LPA in a sample containing the full range of possible PCL-R scores. Only one highly psychopathic subtype emerged, along with two moderate counterparts, and in general, none of the profiles predicted treatment response over and above PCL-R Total Score. These profiles were a replication of an earlier study (Neumann et al., 2016), and since our study, other studies have (again) replicated these three (moderate to strong) psychopathic subtypes in different/various samples containing the full range of scores (Krstic et al, 2018; Lehmann et al., 2019). Apparently, the three profiles are quite a robust finding. For my purpose, it is informative that the two highly psychopathic subtypes that were found in the samples of exclusively high-scoring offenders/patients, were not replicated in any of these studies. Apparently, the two patterns of scores constituting the manipulative and aggressive subtypes respectively, are not robust when tested in the PCL-R studies using the comprehensive samples. In other words, high scores on the PCL-R create more diverse patterns, not just manipulative and aggressive variants. This leads to the conclusion that for my purpose, psychopathic profiles based on the PCL-R do not appear to be a useful avenue, at least not beyond considerations of general severity.

Are There Specific Patient Characteristics That May Be Useful in Tailoring Treatment?
Chapter 3 demonstrated that there is at least one meaningful difference in *outcome* in the treatment of psychopathic patients: about half of the prototypical group dropped out, while the other half not only finished treatment in a regular manner, but also profited equally from treatment as the moderate 'sociopathic' group did, both in terms

of treatment length and recidivism. Therefore, in Chapter 4 we undertook a closer inspection of drop-out. Using interviews with expert clinicians, we explored which factors appeared to make the difference between drop-out and regular treatment completion. Additionally, in our fourth study we chose a different strategy. Instead of focusing on patients (a person-centered approach), we used a variable-centered approach and focused on the PCL-R facets. By using an instrument that portends to tap the full spectrum of psychopathology (MMPI-2-RF), we aimed to amplify our exploration of the PCL-R facets. In our view, this was another way of elucidating the impact of the facets, and to generate hypotheses about possible personalized treatment needs and responsivity issues associated with high scores on specific facets.

Table 1 summarizes patient characteristics from Chapters 4 and 5 that may influence the treatability of highly psychopathic patients. From Chapter 4, only the features that were 'generally' or 'typically' found in the interviews are included (i.e., mentioned in all interviews or all-but-one, and mentioned in more than half of the interviews, respectively). From the viewpoint of the RNR-model, these characteristics include dynamic risk factors as well as responsivity issues. To be able to use these patient characteristics to inform and adapt treatment in an evidence-based manner, further (quantitive) research is needed to confirm their relevance. Moreover, if they are indeed replicated as treatment-relevant factors, there are two possible strategies to reduce drop-out and improve treatability: dynamic factors need to be treated, whereas static factors need a kind of 'work-around' strategy.

When taking a closer look at the characteristics in Table 1 several suggestions can be done. First, high levels of impulsivity and lack of a supportive network are risk factors from the Central Eight, (ie., antisocial personality pattern, antisocial associates, family/marital circumstances). These factors are relevant for any offender and need to be treated 'in their own right' to reduce recidivism after treatment, according to the Need principle. In the case of psychopathic offenders, according to the experts in Chapter 4,

Table 1 *Patient Characteristics Possibly Associated With Drop-out*

Patient characteristics	
Associated with Facet 1	Some minimal ability to accept authority Some minimal degree of trustworthiness
Associated with Facet 2	Degree of: • emotional connection • responsibility for own behavior • mistrust/suspiciousness
Associated with Facet 3	Degree of impulsivity
Not measured by the PCL-R	Age (i.e., older means better prognosis) Degree of motivation Quality of social network and its support for treatment Presence of comorbid borderline traits

they are also deemed to be related to drop-out *during* treatment. This, in a sense, provides an extra 'motive' to carefully assess and treat these factors in psychopathic patients.

Second, all other characteristics in Table 1 appear to be responsivity issues, i.e., factors not related to recidivism but important to take into account to optimize the response to treatment. Age and presence of comorbid borderline personality traits are static factors, they cannot be changed or influenced. If they are replicated as factors that do indeed affect treatability, they may be used to alert clinicians to (for example) a relatively better prognosis (i.e., in older patients with comorbid BPD traits). Motivation is a responsivity issue that again concerns all offenders, but according to the experts in Chapter 4 is especially salient in psychopathic ones. Motivation can be treated with (for example) Motivational Interviewing, which may also encourage patients to take some responsibility for their criminal behavior. The remaining characteristics, all related to the Interpersonal and Affective Facets, are presumed to be (largely) static traits, requiring the above-mentioned 'work-around' strategies. For example, when the ability to form a therapeutic relationship is limited due to defective emotional functioning and a general suspiciousness, treatment may have to depend much more on establishing a framework with clear rules, concrete expectations regarding desirable behavior, and predictable consequences of treatment-interfering behavior. Such a framework may also inspire a minimum degree of trustworthiness on the part of the patient. Finally, difficulties with accepting authority may for example be mitigated by leaving room for individual choices in the day-to-day treatment program.

What Characteristics of the Team and the Treatment May Improve Treatability?

In addition to elucidating treatment-relevant patient characteristics, we also tapped the experts (Chapter 4) for hypotheses about which general aspects of treatment and the team that delivers it need special attention with highly psychopathic patients. Strictly speaking, these suggestions are not directly related to the central theme of this thesis (personalized treatment). However, they may contribute to developing specialized treatment for the subgroup of offenders that scores high on psychopathy. Again limiting myself to the characteristics that were 'generally' or 'typically' mentioned in the interviews, level of expertise and sufficient emotional distance were deemed crucial. Perhaps treating psychopathic patients should be seen as a kind of medical specialty, which requires specialist knowledge, skills, and experience. Mere declarative knowledge about psychopathy is not enough. It seems imperative to have firsthand experience of how truly charming a psychopathic patient can appear, and how good he is at lying. How he can vehemently deny any wrongdoing, making the therapist doubt herself, despite all evidence to the contrary. How he can one day give the therapist the impression of attachment and loyalty, while the next day, having been caught with drugs by his supervisor, he will just as easily betray her and claim they were given to him by the therapist. Our participants suggest it is essential that treatment providers understand what it means that psychopathic patients really are (at least partially) *defective* in the area of human relationships; not just as an intellectual given, but also in practice. For one must then learn not to take ups as well as downs personally.

Regarding the treatment itself, having a coherent underlying treatment philosophy, and an explicit strategy for reacting to treatment interfering behavior were also considered important factors of success. Experienced clinicians felt that a relatively limited inpatient phase worked best, with an extended phase outside the clinic, with gradual, stepwise exposure to situations associated with risk for recidivism. Additionally, they reported that being clear about expectations worked well; allowing some room for mistakes, but with immediate and predictable limit setting. Apparently, participants did not expect much of a therapeutic effect from the inpatient phase of treatment. Possibly this perception is related to pessimism about changing psychopathic personality traits on the one hand, and the subversive role they may play in patient groups on the other hand. Individually tailored behavioral training with gradual exposure to risk then seems a logical treatment strategy. This notion is not really new; I remind the reader of Cleckley's (1941/1976) program of rewarding socially acceptable conduct and punishing irresponsible and destructive behavior (as discussed in the Introduction). A relatively recent law in the Netherlands can support such a strategy. The so-called '*Wet Langdurig Toezicht*' allows for an unlimited period of forensic supervision by probation officers, which could be combined with a relatively limited TBS inpatient treatment phase.

In Conclusion: Psychopathy Through the Lens of the RNR-Model

In my view, applying the RNR-model to the treatment of psychopathy leads to the following conclusions. With respect to the Risk principle: as discussed in the introduction of this thesis, previous research has shown that psychopathy is a high-risk condition, therefore requiring intensive treatment. The research in this thesis mainly concerns the Need and Responsivity principles. TBS treatment focuses on risk reduction, in line with the Need principle. The treatment results in Chapter 3, as well as earlier treatment studies by Wong and colleagues, as discussed in the Introduction (Olver & Wong, 2009; Lewis, Olver, & Wong, 2013), showed that psychopathic patients that do not drop out may profit from regular forensic psychiatric treatment. Nevertheless, high drop-out rates (nearly 50% in our study) remain a serious impediment. We found several individual characteristics that possibly influence the treatability of psychopathic offenders. These characteristics, summarized in Table 1, include several responsivity factors as well as two risk factors. Risk factors are, of course, in general related to recidivism for all offenders. According to our experts, the two risk factors in Table 1 may *also* interfere with treatment itself, leading to drop-out. However, as was discussed above, most of the individual characteristics that were found, are responsivity issues. Hence, it is the Responsivity principle that appears to be crucial to improving treatment/treatability of psychopathic offenders.

Future Research

Much work still needs to be done before we will have a specialized evidence-based treatment manual for psychopathic patients. First, the patient characteristics that were found in our exploratory studies need to be replicated as factors that do indeed moderate treatability, for male as well as female patients. Next, possible strategies to treat or manage these characteristics need to be designed. In working out these

strategies, it appears highly advisable to also pay attention to the specific knowledge, skills, and experience treatment providers need to possess, to be able to carry out these strategies. For example, to be able to keep sufficient emotional distance, supervision or case consultation may be vital. Finally, these treatment strategies must be tested in, for example, series of single case experiments.

An important hurdle to overcome is measuring relevant patient characteristics over time. The PCL-R is a static instrument, as explained in the Introduction, because the scoring is based on life-long patterns as opposed to, for example, the past month or year. It cannot be used to track possible treatment gains. Instruments for risk assessment are also unsuitable, as they are not constructed to measure the specific responsivity issues associated with psychopathy. One solution could be to search for or develop instruments that measure specific traits. Another option is to use the Comprehensive Assessment of Psychopathic Personality (CAPP; Cooke, Hart, Logan, & Michie, 2012; Cooke & Logan, 2018). The CAPP was purposefully developed as a dynamic instrument, potentially capable of detecting change in psychopathic traits and symptoms over time. Patients are scored, using a seven-point scale, on 33 symptoms in 6 domains, named 'Attachment,' 'Behavioral,' 'Cognitive,' 'Dominance,' 'Emotional,' and 'Self.' Scoring starts with a review of the institutional file, followed by a semi-structured clinical interview focusing on a specific time frame (e.g., the past 6-12 months), and/or an Informant Rating Form. Preliminary psychometric evaluations of the CAPP appear promising (see Cooke & Logan, 2018, for a review). There is another advantage to using such an instrument. The primary focus in forensic psychiatry is on reducing criminal behavior. However, this does not mean that forensic treatment may not also *eventually* modify the core personality traits of psychopathy, even if they are currently believed to be (very) resistant, if not impossible to change. I am not aware of any recent studies that have investigated this issue. Repeatedly scoring an instrument such as the CAPP during the process of treatment may also be used to address this question.

There is a final source that I feel needs to be consulted: psychopathic patients themselves. Of course, psychopathic patients are not YAVIS clients (young, attractive, verbal, intelligent, successful). They are not very introspective, often suspicious, and primarily tend to have their own best interests at heart. However, they can also be surprisingly observant about what goes on in treatment, among their fellow patients, and in the institute in general. My personal hypothesis is that this is a survival skill. Psychopathic patients inhabit a hostile, criminal world. This requires a constant alertness. In my opinion we should use this quality in research, to our and *their* advantage.

References

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Author.

Andrews, D. A., & Bonta, J. (2017). *The Psychology of Criminal Conduct* (6th ed.). LexisNexis/Anderson.

Capote, T. (1966). *In cold blood*. Penguin Books; reissued in 2009.

Cleckley, H. (1976). *The mask of sanity. An attempt to reinterpret the so-called psychopathic personality* (5th ed.). Mosby. (Original work published 1941)

Cooke, D. J., Hart, S. D., Logan, C., & Michie, C. (2012). Explicating the construct of psychopathy: Development and validation of a conceptual model, the Comprehensive Assessment of Psychopathic Personality (CAPP). *International Journal of Forensic Mental Health, 11*(4), 242-252. https://doi.org/10.1080/14999013.2012.746759

Cooke, D. J., & Logan, C. (2018). Capturing psychopathic personality: Penetrating the mask of sanity through clinical interview. In C. J. Patrick (Ed.), *Handbook of psychopathy* (2nd ed., pp. 189-210). The Guilford Press.

Cunliffe, T. B., Gacono, C. B., & Smith, J. M. (2021). Appendix B: Descriptions of Eight Select Psychopathic Traits Evident in Women. In J. M. Smith, C. B. Gacono, & T. B. Cunliffe (Eds.), *Understanding female offenders: Psychopathy, criminal behavior, assessment, and treatment*. Academic Press.

Forouzan, E., & Cooke, D. J. (2005). Figuring out la femme fatale: Conceptual and assessment issues concerning psychopathy in females. *Behavioral Sciences & the Law, 23*(6), 765-778. https://doi.org/10.1002/bsl.669

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

Hill, C. E. (Ed.). (2012). *Consensual qualitative research: A practical resource for investigating social science phenomena*. American Psychological Association.

Hill, C. E., Knox, S., Thompson, B. J., Williams, E. N., Hess, S. A., & Ladany, N. (2005). Consensual qualitative research: An update. *Journal of Counseling Psychology, 52*(2), 196-205. https://doi.org/10.1037/0022-0167.52.2.196

Hill, C. E., Thompson, B. J., & Williams, E. N. (1997). A guide to conducting consensual qualitative research. *The Counseling Psychologist, 25*(4), 517-572. https://doi.org/10.1177/0011000097254001

Hopwood, C. J., Bagby, R. M., Gralnick, T., Ro, E., Ruggero, C., Mullins-Sweatt, S., Kotov, R., Bach, B., Cicero, D. C., Krueger, R. F., Patrick, C. J., Chmielewski, M., DeYoung, C. G., Docherty, A. R., Eaton, N. R., Forbush, K. T., Ivanova, M. Y., Latzman, R. D., Pincus, A. L., Samuel, D. B., Waugh, M. H., Wright, A. G. C., & Zimmermann, J. (2020). Integrating psychotherapy with the hierarchical taxonomy of psychopathology (HiTOP). *Journal of Psychotherapy Integration, 30*(4), 477-497. https://doi.org/10.1037/INT0000156

Krstic, S., Neumann, C. S., Roy, S., Robertson, C., 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. J. B., 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

Lewis, K., Olver, M. E., & Wong, S. C. (2013). The Violence Risk Scale: Predictive validity and linking changes in risk with violent recidivism in a sample of high-risk offenders with psychopathic traits. *Assessment, 20*(2), 150-164. https://doi.org/10.1177/1073191112441242

McGuire, J. (2015). 'What works' to reduce re-offending: 18 years on. In: L. A. Craig, L. Dixon, & T. A. Gannon (Eds.), *What works in offender rehabilitation: An evidence-based approach to assessment and treatment* (pp. 20-49). Wiley-Blackwell.

Mokros, A., Hare, R. D., Neumann, C. S., Pekka, S., Habermeyer, E., and Nitschke, J. (2015). Variants of psychopathy in adult male offenders: A latent profile analysis. *Journal of Abnormal Psychology, 124*(2), 372–386. https://doi.org/10.1037/abn0000042

Morrissey, C. (2003). The use of the PCL-R in forensic populations with learning disability. *British Journal of Forensic Practice, 5*(1), 20-24. https://doi.org/10.1108/14636646200300003

Morrissey, C. (2007). *Guidelines for assessing psychopathy in offenders with intellectual disabilities using the PCL-R and PCL:SV*. Retrieved November 15, 2020, from https://www.researchgate.net/publication/257890698_Guidelines_for_use_of_the_PCL-R_and_SV_in_adults_with_ID

Morrissey, C., Hogue, T., Mooney, P., Lindsay, W. R., Steptoe, L., Taylor, J., et al., (2005). Applicability, reliability, and validity of the Psychopathy Checklist-Revised in offenders with intellectual disabilities: Some initial findings. *International Journal of Forensic Mental Health, 4*(2), 207-220. https://doi.org/10.1080/14999013.2005.10471225

Morrissey, C., Hogue, T., Mooney, P., Allen, C. & Johnston, S., Hollin, C., et al. (2007a). The predictive validity of the PCL-R in offenders with intellectual disability in a high secure hospital setting: Institutional aggression. *Journal of Forensic Psychiatry and Psychology, 18*(1), 1-15. https://doi.org/10.1080/08990220601116345

Morrissey, C., Mooney, P., Hogue, T., Lindsay, W. R., & Taylor, J. (2007b). Predictive validity of psychopathy in offenders with intellectual disability in high security: Treatment progress. *Journal of Intellectual and Developmental Disability, 32*(2), 125-134. https://doi.org/10.1080/13668250701383116

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* (pp.14-31). Routledge.

Ng, M. Y., and Weisz, J. R. (2016). Annual research review: Building a science of personalized intervention for youth mental health. *Journal of Child Psychology and Psychiatry, 57*(3), 216-236. https://doi.org/10.1111/jcpp.12470

Nicholls, T. L., & Petrila, J. (2005). Gender and psychopathy: An overview of important issues and introduction to the special issue. *Behavioral Sciences & the Law*, *23*(6), 729-741. https://doi.org/10.1002/bsl.677

Olver, M. E., & Wong, S. C. (2009). Therapeutic responses of psychopathic sexual offenders: Treatment attrition, therapeutic change, and long-term recidivism. *Journal of Consulting and Clinical Psychology*, *77*(2), 328-336. https://doi.org/10.1037/a0015001

Olver, M. E., Sewall, L. A., Sarty, G. E., Lewis, K., and Wong, S. C. P. (2015). 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. https://doi.org/10.1037/abn0000038

Wong, S. C. P., & Hare, R. D. (2005). *Guidelines for a psychopathy treatment program*. Multi-Health Systems.

Wong, S. C., & Olver, M. E. (2015). Risk reduction treatment of psychopathy and applications to mentally disordered offenders. *CNS spectrums*, *20*(3), 303-310. https://doi.org/10.1017/S1092852915000322

Wynn, R., Høiseth, M. H., & Pettersen, G. (2012). Psychopathy in women: theoretical and clinical perspectives. *International Journal of Women's Health*, *4*, 257-263. http://dx.doi.org/10.2147/IJWH.S25518

Individual Differences in the Treatment of Psychopathic Offenders

Psychopathy is a highly relevant clinical condition in forensic psychiatry. Perhaps best understood as a personality disorder, its prevalence in the criminal justice system is quite high with estimates ranging between 15 and 25%, and it is associated with high risk of recidivism. Previous studies have shown that psychopathic patients respond poorly to treatment, with high rates of treatment interfering behavior, aggressive incidents, and drop-out. Nevertheless, little has been done to develop specialized treatment programs. A relevant distinction regarding the goal of treatment, is whether the main focus is fundamentally changing psychopathic personality or 'merely' reducing risk factors for criminal behavior. In forensic psychiatry, for the sake of society, reducing the risk of recidivism is a primary goal.

Correctional treatment in general should adhere to the Risk, Need, and Responsivity principles (RNR-model). These principles state that forensic treatment is most effective (1) when it is matched to level of risk (i.e., higher risk implying more intensive treatment), (2) when it targets risk factors associated with reoffending, and (3) when treatment is adapted to the learning style of the offender (responsivity). To date, two studies have shown that treatment developed for high-risk offenders in general (according to the RNR-model) may lead to reduction of recidivism for those psychopathic patients who did not drop out. Apparently, treatment can be effective for at least some psychopathic offenders. An important follow-up question is whether forensic treatment can be adapted to personal treatment-relevant characteristics of psychopathic patients. Presumably, not all psychopathic patients have identical treatment needs. What, for example, distinguished the completers from the non-completers, and those who recidivated from those who did not? Are some psychopathic subtypes more treatable than others? These questions inspired the present thesis. The central question is: Can individual differences in psychopathic offenders be used systematically to inform and improve their forensic psychiatric treatment with the goal of reducing recidivism? In all studies we used the Psychopathy Checklist-Revised (PCL-R) to measure the level of psychopathy. Furthermore, except in the first study, all samples consisted of patients with a so-called TBS treatment order of indefinite length (*Ter beschikking stelling;* at the disposal of the state, i.e., mandatory treatment for violent and/or sexually violent offenders).

In Chapter 2 two famous cases were presented, the multiple homicide perpetrators from the non-fiction book *In Cold Blood* by Truman Capote. Capote described the lives of Dick and Perry in such detail that it proved possible to score their respective PCL-R items. The two men had very different backgrounds and personalities, and therefore seemed useful in illustrating salient individual differences in psychopathy. The offenders accrued an identically high PCL-R total score (31), but the underlying facet profiles showed meaningful differences. Dick scored higher on glib and deceitful behavior (PCL-R Facet 1), whereas Perry scored higher on antisocial and aggressive behavior (PCL-R Facet 4). Remarkably, these results mirrored the findings of two large scale empirical studies that investigated subtypes in exclusively high-scoring samples of male offenders. Both studies found two subtypes that can be labeled as a manipulative and an aggressive variant. These two subtypes seem to differ primarily in their preferred method of achieving

their (antisocial) goals; one preferring deception and manipulation, the other violent and intimidating behavior. Dick and Perry seemed to embody exactly these different strategies in their criminal behavior, and may have had different treatment needs (but both men were hanged instead).

Chapter 3 employed a quantitative approach to seek individual differences in a sample containing not just high scoring male TBS patients, but these offenders covered the full range of PCL-R scores, hence allowing for greater individual variation. My goals were to investigate whether/which subtypes would emerge, and to what extent these subtypes were clinically useful. We did not replicate the high-scoring manipulative and aggressive variants. Instead, we found two moderately psychopathic profiles (called 'sociopathic' and 'callous-conning', respectively) and only one high-scoring 'prototypical' one. To test the clinical utility of the profiles, we investigated how these subtypes were associated with various treatment variables. The most salient finding was the high drop-out rate in the prototypical group. However, prototypical patients who remained in treatment proceeded through the treatment phases in much the same way as their less severe sociopathic counterparts, and showed comparable recidivism rates as well (i.e., similar benefit from treatment). However, the typology was significantly related to increasing PCL-R total scores. Hence, differential treatment outcomes may have been simply related to the severity of the psychopathic traits in general, rather than to certain subtypes.

To take a step back, Chapter 4 sought to capture the knowledge of experts on individual differences in the treatability of psychopathic offenders. Following a qualitative research approach, we focused on drop-out of the prototypical psychopathic group, and explored what, according to expert clinicians, might distinguish completers from non-completers. Specifically, 11 expert clinicians were interviewed about two specific cases with a PCL-R score of 28 or higher, one considered a treatment failure, the other a treatment success. The interview covered patient characteristics, treatment (provider) characteristics and other factors the participants deemed associated with drop-out/successful completion. High scores on glib and deceitful interpersonal style (Facet 1) and defective affective experience (Facet 2) were thought to impede treatment retention, particularly by their negative impact on motivation and the therapeutic relationship, as was, to a lesser extent, impulsivity (Facet 3). Additionally, several factors not related to psychopathy were mentioned by the informants: Patients with a relatively prosocial, supportive network were deemed to fare better, as were patients with comorbid borderline traits, and/or older patients (presumably because of physical deterioration and/or having "calmed down" over the course of life). With respect to the treatment framework, the following characteristics were volunteered as important: a treatment program that 1) stipulates clear and concrete goals and expectations, including a strategy for dealing with treatment interfering behavior, 2) is offered by a knowledgeable, stable team with sufficient emotional distance, and 3) provides a long and gradual resocialization trajectory. A relatively recent law in the Netherlands, the '*Wet Langdurig Toezicht*', makes such a strategy possible, allowing for an unlimited period of forensic supervision by probation officers, which could be combined with a relatively limited TBS inpatient treatment phase.

Chapter 5 switched from a person-centered approach to a variable-centered approach. We investigated associations between the PCL-R Four-Facet model and the MMPI-2-RF scales, using an exploratory analytic strategy to predict PCL-R scores from the MMPI-2-RF. The guiding objective was to elucidate the nature of the four facets, possibly leading to specific treatment needs associated with specific facets. Most of the scales that were found to be predictive of the PCL-R were quite straightforward, as in fully in line with conceptually expected associations. From the viewpoint of individual differences, the most interesting finding was that the Affective Facet (Facet 2) alone was predicted by Ideas of Persecution (RC6). This scale refers to beliefs that others seek to harm you and is associated with suspicion, alienation, and blaming others for your difficulties. These characteristics are not measured by the PCL-R, and it may be worthwhile to develop specialized treatment strategies to address this issue.

Chapter 6 pursued yet another aspect of individual differences. We examined the question of measurement invariance (MI) of the PCL-R with respect to gender. Establishing MI is a precondition for the comparability of male and female scores. If MI is violated, differences in the item scores between males and females do not necessarily indicate that there are "real" differences in the underlying pathology. Our results indicated that the PCL-R in its current form does not attain full MI. Four items showed threshold-biases and particularly the Social Deviance factor (Factor 2) was gender biased. Apparently, psychopathic females differ from men not so much in core psychopathic personality traits, but rather in the way they express their impulsive, irresponsible lifestyle and antisocial behavior. For example, the use of physical force or violence to achieve a desired outcome may be considerably less feasible and effective for women, whereas the manipulative use of flirtation, intimacy or sexual favors may work better. To improve assessment of psychopathy in females, in our view it is most feasible to develop supplementary guidelines to the existing manual. More elaborate specification of what type of behavior in females is needed for a score of 1 (*may apply or partly applies*), or 2 (*definitely applies*), may improve assessment. Better fidelity in assessment also is a prerequisite for studying treatment responsivity and possible individual differences in females.

In Chapter 7 the present body of work is discussed. A summary of the key findings related to the central theme of these thesis is presented, and methodological considerations are addressed, before turning to the implications for treatment and future research. The RNR-model is used as a framework to discuss treatment implications. Previously, it was found that psychopathic patients are high-risk offenders (Risk principle). The research in this thesis mainly concerns the Need and Responsivity principles. With regard to the Need principle: psychopathic patients that do not drop out appear to be able to profit from regular forensic psychiatric treatment. Nevertheless, high drop-out rates remain a serious impediment. Some dynamic risk factors for recidivism may (also) interfere with treatment, such as severe impulsivity and lack of a supportive network. With regard to the Responsivity principle: in the case of high scoring offenders, PCL-R-based subtypes do not appear to be useful in tailoring treatment to individual cases. However, certain patient characteristics found in Chapters 4 and 5 possibly are detrimental to treatment,

such as high levels of suspiciousness, inability to form even a minimal emotional connection with treatment providers, and/or complete lack of motivation. Hence, it is the Responsivity principle that appears to be crucial to improving treatment/treatability of psychopathic offenders.

Recommendations for future research include replicating the individual characteristics as factors that do indeed moderate treatability, and subsequently designing and evaluating strategies to assess, treat, and/or manage these characteristics. Finally, there is another source that may be consulted: psychopathic patients themselves. These patients can be surprisingly observant, possibly because they inhabit a hostile world, requiring constant vigilance. This quality may be used in research, ultimately to their advantage.

Individuele verschillen in de behandeling van psychopathische daders

Psychopathie is een zeer relevante klinische aandoening in de forensische psychiatrie, die misschien het best begrepen kan worden als een persoonlijkheidsstoornis. De prevalentie in het strafrechtsysteem is vrij hoog, met schattingen variërend tussen 15 en 25%. Psychopathie hangt samen met een hoog risico op recidive. Eerdere studies hebben bovendien aangetoond dat psychopathische patiënten slecht reageren op behandeling, met veel therapie-interfererend gedrag, agressieve incidenten en *drop-out*. Toch is er tot op heden weinig voortgang geboekt in het ontwikkelen van gespecialiseerde behandel-programma's voor deze patiënten. Een essentieel onderscheid in de behandeling is of de belangrijkste focus ligt op het fundamenteel veranderen van de psychopathische persoonlijkheid of 'slechts' op het verminderen van risicofactoren voor crimineel gedrag. In de forensische psychiatrie is het terugdringen van het recidiverisico, in het belang van de maatschappij, het primaire doel.

Een forensische behandeling moet in het algemeen voldoen aan de principes van het zogenaamde RNR-model (*Risk, Need, Responsivity*). Deze principes stellen dat forensische behandeling het meest effectief is (1) wanneer deze is afgestemd op de hoogte van het recidiverisico (waarbij een hoger risico een intensievere behandeling impliceert), (2) wanneer de behandeling gericht is op risicofactoren die samenhangen met recidive, en (3) wanneer de behandeling is aangepast aan de leerstijl van de dader. Twee eerdere studies hebben aangetoond dat behandeling die is ontwikkeld voor hoog-risico daders in het algemeen (volgens het RNR-model) kan leiden tot vermindering van de recidive bij psychopathische patiënten die deze behandeling voltooiden. Blijkbaar kan behandeling effectief zijn voor tenminste een deel van de psychopathische daders. Een belangrijke vervolgvraag is of een dergelijke 'reguliere' forensische behandeling voor deze groep kan worden verbeterd door de behandeling aan te passen aan specifieke persoonlijke kenmerken van deze patiënten. Vermoedelijk hebben niet alle psychopathische patiënten dezelfde probleemgebieden. Zijn sommige psychopathische subtypes beter te behandelen dan andere? Wat onderscheidt bijvoorbeeld de *drop-outs* van degenen die de behandeling wel voltooiden? En hoe verschillen degenen die recidiveren van degenen die dat niet doen? Deze vragen vormden de inspiratie voor dit proefschrift. De centrale vraag is: Kunnen individuele verschillen bij psychopathische daders systematisch worden ingezet om hun forensische behandeling te verbeteren met als primair doel om de recidive te verminderen? In ons onderzoek naar deze vraag hebben we gebruik gemaakt van de Psychopathie Checklist-Revised (PCL-R) om het niveau van psychopathie te meten. Verder bestonden, behalve in hoofdstuk 2, alle steekproeven uit patiënten aan wie een ongemaximeerde maatregel van terbeschikkingstelling met dwangverpleging was opgelegd (TBS).

In hoofdstuk 2 worden twee bekende daders gepresenteerd, de meervoudige moordenaars uit het *true crime* boek *In Cold Blood* van Truman Capote. Capote beschreef het leven van Dick en Perry zo gedetailleerd dat het mogelijk bleek om voor beide de PCL-R-items te scoren. De twee mannen hadden een zeer verschillende achtergrond en persoonlijkheid, en leken daarom bij uitstek geschikt om individuele verschillen te illustreren. Zij behaalden een identieke hoge PCL-R totaalscore (31), maar de scores

op de onderliggende facetten vertoonden duidelijke verschillen. Dick scoorde hoger op charmerend, manipulatief en bedrieglijk gedrag (PCL-R Facet 1), terwijl Perry hoger scoorde op antisociaal en agressief gedrag (PCL-R Facet 4). Opmerkelijk is dat deze resultaten overeenkomen met de bevindingen van twee grootschalige empirische studies die subtypes onderzochten in steekproeven van mannelijke delinquenten met een hoge mate van psychopathie. Beide studies vonden twee subtypes: een manipulatieve en een agressieve variant. Deze twee subtypes lijken vooral te verschillen in de methode die zij primair gebruiken om hun (antisociale) doelen te bereiken: de een bij voorkeur door middel van misleiding en manipulatie, de ander met behulp van gewelddadig en intimiderend gedrag. Dick en Perry leken precies deze verschillende strategieën te illustreren in hun criminele gedrag. In een TBS-behandeling had dit waarschijnlijk geleid tot verschillende behandeldoelen (maar beide mannen werden in plaats daarvan opgehangen).

In hoofdstuk 3 wordt gebruik gemaakt van een kwantitatieve benadering om individuele verschillen te onderzoeken, in een steekproef van mannelijke TBS-patiënten met elke mogelijke PCL-R score, niet uitsluitend de hoge. Hierdoor was sprake van een grotere variatie in scores. Mijn doelen waren om te onderzoeken welke subtypes we in een dergelijke groep zouden aantreffen, en in hoeverre deze subtypes bruikbaar zijn in de klinische praktijk. De hoog-scorende manipulatieve en agressieve varianten werden niet gerepliceerd. In plaats daarvan vonden we twee matig psychopathische profielen (respectievelijk 'sociopathisch' en 'ongevoelig' genoemd) en slechts één hoog-scorend 'prototypisch' profiel. Om de klinische bruikbaarheid van de profielen te testen, onderzochten we hoe deze subtypes samenhingen met verschillende behandel-variabelen. De meest opvallende bevinding was (opnieuw) de hoge *drop-out* in de prototypische groep. Echter, prototypische psychopathische patiënten die de behandeling wel voltooiden, doorliepen de behandelfasen op vrijwel dezelfde manier als hun minder ernstige sociopathische medepatiënten, en vertoonden ook vergelijkbare recidivecijfers (dat wil zeggen: zij profiteerden op vergelijkbare wijze van de behandeling). De typologie op basis van de PCL-R scores was echter significant gerelateerd aan de hoogte van de totaalscores. Dat betekent dat de behandelresultaten eenvoudigweg verklaard kunnen worden door de mate van psychopathie, in plaats van door bepaalde subtypes.

Omdat de typologie vooralsnog niets overleverde voor de klinische praktijk zijn we overgestapt op een meer exploratieve aanpak. In hoofdstuk 4 gebruikten we een kwalitatieve onderzoeksmethode om kennis van experts vast te leggen over individuele verschillen in behandelbaarheid van mannelijke psychopathische TBS-ers. Daarbij hebben we ons gericht op de hoge *drop-out* van de prototypische psychopathische groep en geëxploreerd wat volgens deskundige clinici belangrijke verschillen zijn tussen de patiënten die de behandeling wel en niet voltooiden. Elf deskundigen werden geïnterviewd over twee specifieke gevallen met een PCL-R-score van 28 of hoger, waarvan de behandeling bij de een beschouwd werd als mislukt, bij de ander juist als een succes. Het interview had betrekking op patiëntkenmerken, kenmerken van de behandelsetting en andere factoren die volgens de deelnemers verband hielden met uitval versus

succesvolle afronding. Hoge scores op een charmerende en bedrieglijke interpersoonlijke stijl (Facet 1) en een gebrekkig gevoelsleven (Facet 2) belemmerden volgens de deskundigen de behandeling, met name door hun negatieve invloed op de motivatie en de therapeutische relatie. Impulsiviteit (Facet 3) speelde volgens hen eveneens een rol, maar minder sterk. Daarnaast werden door de informanten verschillende factoren genoemd die niet gerelateerd waren aan psychopathie: Patiënten met een relatief prosociaal, ondersteunend netwerk werden geacht het beter te doen, evenals patiënten met comorbide borderline-kenmerken en/of oudere patiënten (dit laatste vermoedelijk vanwege fysieke achteruitgang en/of "in rustiger vaarwater komen" in de loop van het leven). Met betrekking tot de behandelsetting werden de volgende kenmerken als belangrijk aangemerkt: een behandelprogramma dat 1) duidelijke en concrete doelen en verwachtingen stelt, inclusief een strategie om met regel-overtredend gedrag om te gaan, 2) dat wordt aangeboden door een deskundig, stabiel team met voldoende emotionele distantie, en 3) een lang en geleidelijk resocialisatietraject biedt. Een relatief recente wet in Nederland, de Wet Langdurig Toezicht, maakt een dergelijke strategie mogelijk, waarbij een onbeperkte periode van forensisch toezicht door reclasseringsambtenaren mogelijk is, eventueel gecombineerd met een relatief beperkte klinische behandelfase in een TBS-kliniek.

In hoofdstuk 5 schakelden we van een persoonsgerichte benadering over op een benadering gebaseerd op testvariabelen. We onderzochten de samenhang tussen het vier-facetten-model van de PCL-R en de schalen van de MMPI-2-RF, een instrument dat op veel uitgebreidere wijze het persoonlijkheidsfunctioneren meet. Hierbij werd gebruik gemaakt van een exploratieve analytische strategie, waarbij PCL-R-scores werden voorspeld vanuit de MMPI-2-RF. Het doel was om met behulp van de MMPI-2-RF de aard van de vier facetten te verhelderen, mogelijk leidend tot specifieke behandeldoelen die verband houden met specifieke facetten. De meeste schalen die voorspellend bleken te zijn voor de PCL-R waren in lijn met de conceptueel te verwachten associaties. Vanuit het oogpunt van individuele verschillen was de meest interessante bevinding dat het affectieve facet (Facet 2) werd voorspeld door *Ideas of Persecution* (RC6). Deze schaal meet overtuigingen dat anderen je kwaad willen doen en is geassocieerd met achterdocht, vervreemding en het anderen de schuld geven van moeilijkheden. Deze kenmerken worden niet gemeten door de PCL-R. Mogelijk is het de moeite waard om gespecialiseerde behandelstrategieën te ontwikkelen die expliciet rekening houden met achterdocht bij psychopathische patiënten.

De studie in hoofdstuk 6 gaat in op nog een ander aspect van individuele verschillen. In deze studie onderzochten we de kwestie van meetinvariantie (MI) van de PCL-R met betrekking tot sekse. Het vaststellen van MI is een voorwaarde voor de vergelijkbaarheid van de scores van mannen en vrouwen. Als MI wordt geschonden, dan wijzen verschillen in de itemscores tussen mannen en vrouwen er niet noodzakelijkerwijs op dat er daadwerkelijk "echte" verschillen zijn in de onderliggende pathologie. Onze resultaten geven aan dat de PCL-R in zijn huidige vorm geen volledige MI bereikt. Vier items functioneerden problematisch en vooral bij de zogenaamde sociale deviantie factor

(Factor 2) was sprake van gender bias. Blijkbaar verschillen psychopathische vrouwen niet zozeer van mannen in de belangrijkste psychopathische persoonlijkheidskenmerken (Factor 1), maar eerder in de manier waarop ze uiting geven aan een impulsieve, onverantwoordelijke levensstijl en antisociaal gedrag. Zo kan het gebruik van fysiek geweld om een gewenst resultaat te bereiken voor vrouwen aanzienlijk minder haalbaar en effectief zijn, terwijl het manipulatieve gebruik van flirten of seksuele gunsten wellicht beter werkt. Om de beoordeling van psychopathie bij vrouwen te verbeteren, is het naar onze mening het meest effectief om aanvullende richtlijnen op de bestaande handleiding te ontwikkelen. Een meer uitgebreide specificatie van welk type gedrag bij vrouwen nodig is voor een score van 1 (*gedeeltelijk van toepassing*) of 2 (*zeker van toepassing*) kan de beoordeling verbeteren. Een betere betrouwbaarheid van het instrument is ook een voorwaarde voor het bestuderen van mogelijke individuele verschillen in behandelbaarheid bij vrouwen.

In hoofdstuk 7 worden de studies van dit proefschrift in onderlinge samenhang besproken. Een samenvatting van de belangrijkste bevindingen met betrekking tot het centrale thema van dit proefschrift wordt gepresenteerd, en methodologische overwegingen worden behandeld. Het RNR-model wordt gebruikt als raamwerk om de implicaties voor de behandeling te bespreken. Eerder onderzoek toonde aan dat psychopathische patiënten hoog-risico daders zijn (*Risk*-principe). Het onderzoek in dit proefschrift heeft voornamelijk betrekking op de *Need*- en *Responsivity*-principes. Wat betreft het *Need*-principe: psychopathische patiënten die niet uitvallen blijken baat te kunnen hebben bij reguliere forensisch psychiatrische behandeling. Desalniettemin blijft de hoge *drop-out* een ernstige belemmering. Sommige dynamische risicofactoren voor recidive kunnen (ook) leiden tot therapie-interfererend gedrag tijdens de behandeling, zoals ernstige impulsiviteit en het ontbreken van een ondersteunend netwerk. Wat betreft het *Responsivity*-principe: in het geval van hoog-scorende daders lijken op de PCL-R gebaseerde subtypes niet bruikbaar om de behandeling te verbeteren. Echter, bepaalde specifieke patiëntkenmerken die gevonden werden in hoofdstukken 4 en 5 zijn mogelijk wel van invloed op de behandeling, zoals een hoge mate van achterdocht, het onvermogen om zelfs maar een minimale emotionele band aan te gaan met behandelaren en/of een volledig gebrek aan motivatie. Daarom lijkt het *Responsivity*-principe cruciaal voor het verbeteren van de behandeling/behandelbaarheid van psychopathische patiënten.

Aanbevelingen voor toekomstig onderzoek bestaan uit het repliceren van de individuele kenmerken als factoren die inderdaad de behandelbaarheid beïnvloeden, en vervolgens het ontwerpen en evalueren van strategieën om deze kenmerken te meten, te behandelen en/of te managen. Ten slotte is er nog een andere bron die kan worden geraadpleegd: psychopathische patiënten zelf. Deze kunnen verrassend opmerkzaam zijn. Psychopathische daders zijn gewend om in een vijandige, criminele wereld te leven die constante waakzaamheid vereist. Deze opmerkzaamheid kan worden gebruikt in wetenschappelijk onderzoek. Uiteindelijk is dit ook in hun voordeel.

DANKWOORD

Er is een Engels gezegde dat luidt: *It takes a village to raise a child*. Hetzelfde geldt mijns inziens voor het schrijven van een proefschrift. Er zijn zo veel mensen die direct of indirect hebben bijgedragen aan mijn promotietraject. Het begon in de Van der Hoevenkliniek met Daan van Beek, die mij op het idee bracht, en mij in veel wat ik deed aanmoedigde. Ik denk met plezier terug aan onze lunches in Utrecht, waarin nog zoveel meer aan bod kwam dan alleen onderzoek. Het begon ook met Monique Kossen, die mij tijd en mogelijkheden gaf om aan dit enorme project te beginnen. Tijd en mogelijkheden kreeg ik vervolgens ook bij De Waag Nederland, met dank aan Coen van Gestel en Evelien van Dijk, en bij de Oostvaarderskliniek, met dank aan Hendrik Jan van der Lugt. Ik mag ook de patiënten met een hoge mate van psychopathie niet vergeten, die ik in behandeling heb gehad. Jullie waren in zekere zin overgeleverd aan mijn nieuwsgierigheid naar jullie denkwijze en belevingswereld. Velen van jullie waren bereid daar een inkijkje in te bieden, en daar begon mijn zoektocht naar behandelmogelijkheden.

Zonder de onderzoeksafdeling van de Van der Hoevenkliniek had ik het niet gekund, ik denk dan in het bijzonder aan Vivienne de Vogel en Edwin Wever, maar ook aan de andere onderzoekers, diagnostici en stagiaires. Wat een indrukwekkende databestanden hebben jullie de afgelopen 20 jaar opgebouwd, waar ik gebruik van mocht maken. Er zijn daarnaast 11 behandelaren die ik vanwege de anonimiteit niet met naam mag noemen, die hebben deelgenomen aan het kwalitatieve onderzoek in hoofdstuk 4. De interviews met jullie reken ik tot de meest interessante momenten van de afgelopen jaren. Ik kan in dit verband alleen Mieke Breij met naam en toenaam noemen, die bereid was als proefpersoon te dienen voor het interview. Medeauteurs Craig Neumann en Dylan Molenaar waren van onschatbare waarde voor de statistiek, en Kelsey Timmer voor het uitwerken en helpen bij de kwalitatieve analyse van de interviews. Hoeveel vrijdagen hebben wij niet met Wineke Smid doorgebracht, discussiërend over tekstfragmenten? (Vaak gecaterd door Dorien, waarvoor dank!)

Heel dankbaar ben ik ook voor de steun van mijn partner en kinderen. Jullie hebben geaccepteerd dat ik nou eenmaal een beetje een vakidioot ben, die ook haar vrijetijd heeft opgeofferd voor dit project. Jullie leefden mee, ook op de momenten dat het tegenzat, zoals toen ik een volle kop koffie in mijn laptop liet vallen, in de tijd dat mijn documenten nog niet in een *cloud* stonden.

En dan *last but not least*, integendeel, zondermeer het meest belangrijk waren Jan Henk Kamphuis en Wineke Smid. Jaar in jaar uit, altijd constructief kritisch, altijd steunend, altijd inspirerend. Jullie hielpen me denken; het werd er altijd beter van. Ik zal nooit de uren vergeten die we op het Roeterseiland hebben doorgebracht in de kantoren die jij daar in de loop der jaren hebt gehad, Jan Henk.
Ik heb het een groot, groot voorrecht gevonden om door jullie begeleid te worden. Dank daarvoor.

· | -| ⊣ + + → -⟩ · ⟩
oostvaarderskliniek