

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/44736> holds various files of this Leiden University dissertation

**Author:** Spaans, Marleen

**Title:** Personality pathology in a forensic setting : prevalence, assessment, and prognostic value for treatment

**Issue Date:** 2016-12-06

# Personality Pathology in a Forensic Setting

Prevalence, Assessment, and Prognostic Value for Treatment



Marleen Spaans



Dienst Justitiële Inrichtingen  
*Ministerie van Veiligheid en Justitie*

Dit proefschrift is tot stand gekomen met financiering en medewerking van het Nederlands Instituut voor Forensische Psychiatrie en Psychologie.

# **Personality pathology in a forensic setting Prevalence, assessment, and prognostic value for treatment**

**Proefschrift**

ter verkrijging van  
de graad van Doctor aan de Universiteit Leiden,  
op gezag van Rector Magnificus Prof. mr. C.J.J.M. Stolker,  
volgens besluit van het College voor Promoties  
ter verdediging op dinsdag 6 december 2016  
klokke 11.15 uur

door

**Marleen Anna Spaans**

geboren te Bangkok  
op 3 oktober 1981

**Promotores**

Prof. dr. P. Spinhoven

Prof. dr. E. de Beurs

**Co-promotor**

Dr. T. Rinne (*Nederlands Instituut voor Forensische Psychiatrie en Psychologie*)

**Promotiecommissie**

Prof. dr. L. Eurelings-Bontekoe

Prof. dr. J. van der Leun

Prof. dr. R. Vermeiren

Prof. dr. F. Koenraadt (*Universiteit Utrecht*)

Prof. dr. H. van Marle (*Erasmus University Rotterdam*)

Voor mijn ouders



# Table of contents

<b>Chapter One</b>	
Introduction	6
<b>Chapter Two</b>	
Self-reported personality traits in forensic populations: A meta-analysis	16
<b>Chapter Three</b>	
The DAPP-SF as a screener for personality disorder in a forensic setting	36
<b>Chapter Four</b>	
MMPI profiles of males accused of severe crimes, a cluster analysis	54
<b>Chapter Five</b>	
Diagnosis of antisocial personality disorder and criminal responsibility	66
<b>Chapter Six</b>	
Prognostic factors for length of enforced forensic treatment with special focus on personality disorder	78
<b>Chapter Seven</b>	
Discussion	94
<b>References</b>	108
<b>References from the current dissertation</b>	121
<b>Appendix 1</b>	
Search terms for meta-analysis and systematic review	122
<b>Appendix 2</b>	
Flow diagram of search process for meta-analysis and systematic review	123
<b>Dutch summary</b>	124
<b>Acknowledgements</b>	133
<b>Curriculum Vitae</b>	134



Chapter One  
*Introduction*

According to criminal codes from many international jurisdictions, an offender can be considered less or not punishable if he or she commits a crime as the result of a mental disorder that influenced that individual's free will, so that it rendered him or her unable to act differently. In some cases, courts can decide on a mitigated sentence or compulsory treatment for the offender. Article 39 of the Dutch criminal code states: "not punishable is he who commits a criminal act for which he cannot be held responsible because of a defect or disorder of his mental capacities". Under Dutch criminal law, if an offender's mental disorder meant that he or she did not have the opportunity to act differently and the crime was unavoidable, the fact cannot be attributed to him or her and courts should refrain from punishment. Instead of punishment, courts can order enforced treatment, if there is a danger to society of the same crime being committed as a result of the mental disorder's influence on the individual's freedom to act (Mooij, 20014). In the Netherlands, the presence of a personality disorder at the time of the crime can warrant diminished criminal responsibility and/or enforced treatment in a high security hospital (Barendregt, Muller, Nijman, & De Beurs, 2008; De Kogel & Nagtegaal, 2006; Nijman, De Kruyk, & Van Nieuwenhuizen, 2004). The Dutch system of enforced treatment of (personality-)disordered offenders has been under close public and political scrutiny over the last decade, following a number of serious incidents. The number of disordered offenders in treatment hospitals greatly increased in recent years (from 405 individuals in 1990 to over 2100 in 2010). Also, the system has undergone many policy changes, but these changes have been so frequent and followed one another so quickly that many policies were reversed before they could be evaluated (Nagtegaal, Van der Horst, & Schönberger, 2011). According to Nijman and colleagues (2004), offenders with a personality disorder outnumber those with psychotic disorders in treatment hospitals. Therefore, it is essential to be able to make thorough and accurate diagnoses of (the severity of) personality pathology – consisting of both personality disorders and personality characteristics – in psychological and psychiatric assessments within the field of the interface, or common ground, of criminal law and psychology.

## Standardization of psychodiagnostic assessment in the Pieter Baan Center

The Pieter Baan Center (PBC) is the Dutch Ministry of Justice's psychiatric observation hospital that carries out around 220 pre-trial in-patient assessments of suspects of serious crimes per year. All defendants are evaluated during a seven-week period by a multidisciplinary team consisting of a psychiatrist, a psychologist, a forensic social worker, a ward staff member, and a legal expert who supervises the assessment process along with a second senior psychiatrist or psychologist. The forensic social worker investigates the life history and social background of the defendant through interviews with informants such as family members, former teachers, or former employers. The ward staff member has the task of observing and describing the activities and behavior of the defendant during his or her stay on one of the observation hospital's wards. The final product of an assessment in the PBC is a report by the forensic mental health experts concerning each defendant's psychiatric disorders, if any, and degree of criminal responsibility. The final report also contains a recommendation to the court as to whether enforced treatment of the defendant is considered necessary if the defendant is convicted of the charge. The PBC's recommendation is followed by the court in around 86% of all cases (Boonekamp, Barendregt, Spaans, De Beurs, & Rinne, 2008).

Psychodiagnostic assessment is a very important information source during an evaluation in the PBC. Until the end of 2007, the decision as to which assessment instruments to use was determined per case by the examining psychologist and psychiatrist. Their aim was to find an optimal balance between the expertise of each expert, the case-specific details, and the available assessment instruments. A major drawback of such an expert and practice-based approach, however, was that it had the potential to compromise the reproducibility – and especially the comparability – of assessment results. Moreover, this method was relatively vulnerable to thinking and decision-making errors such as confirmation bias (De Ruiter, 2007; Barendregt, Rijnders, & De Ranitz, 2008).

Due to this criticism, the PBC's psychological assessment procedure and test instruments were critically examined and revised in 2007. The revision had many aims, including to achieve greater uniformity and standardization of the psychiatric and psychological assessment process, to promote evidence-based assessment by increasing the use of scientifically sound and reliable instruments, and to establish a new diagnosis protocol based on objective and standardized measurement methods and assessment instruments for a complex population within the field of criminal law otherwise known as a forensic population (Spaans, Barendregt, Muller, Van der Meer, & Rinne, 2014).

A positive and intended result of standardizing the assessment process in the PBC was the gradual creation of a database of defendants' scores on a number of assessment instruments. The new battery of assessment instruments was intended to be dynamic, meaning it could be adjusted at any time. This database made it possible to study the utility of the instruments selected for the assessment protocol, as many of the instruments used or even available in the Netherlands at the time were not specifically designed for forensic population and utility in the PBC's populations was unknown.

The current study investigates the prevalence, assessment and prognostic value for treatment of personality pathology – which within the context of the present study refers to both personality disorders and maladaptive personality traits – in a forensic mental health setting, largely using the database of results on assessment instruments acquired after the revision of the assessment process in the PBC.

## Personality pathology in a forensic mental health setting

Personality pathology is highly prevalent in prison populations. According to a systematic review carried out by Fazel and Danesh in 2002, 65% of the male general prison population and 42% of the female general prison population in western countries has a personality disorder, based on interview methods of assessing these disorders. The authors also found that prisoners are ten times more likely to have antisocial personality disorder than the general population, with a prevalence of 47% in men and of 21% in women. De Ruiter and Greeven (2000), Hildebrand and De Ruiter (2004), and Timmerman and Emmelkamp (2001) found that 80%, 88.3%, and 87% of their forensic samples, respectively, had at least one personality disorder diagnosis. In all three samples the most frequently diagnosed personality disorders originated from the American Psychiatric Association's (APA) Diagnostic Statistical Manual's (DSM) Cluster B: antisocial, borderline, histrionic and narcissistic personality disorders.

Cluster B personality disorders, along with Cluster A, are considered to be more severe than Cluster C (Crawford, Koldobsky, Mulder, & Tyrer, 2011). DSM-IV-TR Cluster B antisocial personality disorder and narcissistic personality disorder are characterized by maladaptive personality traits such as high levels

of aggression or hostility, poor impulse control, sensation seeking, and lack of empathy (American Psychiatric Association, 2000; Hare, 2006; Looper & Paris, 2000). These personality traits are associated with impairments or abnormalities that lead to maladaptive outcomes and are also especially prevalent in forensic populations (Boccaccini, Murrie, Hawes, Simpler, & Johnson, 2010; Cunradi, Todd, Duke, & Ames, 2009; Dolan & Blackburn, 2006; Norlander & Eckhardt, 2005; Shechory, Weiss, & Weinstain, 2011; Walters, 2007).

## Assessment of personality pathology in a forensic mental health setting

Given that previous research has shown high prevalence in forensic populations of personality disorder and personality traits that lead to negative outcomes, the assessment of personality pathology (i.e., both personality *disorders* and *traits*) is a very important part of the legal process. Not only is knowledge about an individual's personality pathology of great importance for the judge and his or her decisions regarding punishment and possible treatment – especially in those jurisdictions in which the presence of a personality disorder warrants diminished criminal responsibility – but it also plays a large role in the treatment process of those offenders.

The APA's guidelines for forensic psychology state that forensic mental health practitioners should "provide opinions and testimony that are sufficiently based upon adequate scientific foundation, and reliable and valid principles and methods that have been applied appropriately to the facts of the case" (American Psychiatric Association, 2013b; p.9). The assessment of personality pathology in the forensic mental health field is, however, more complex than in regular psychological practice. Firstly, most defendants have not requested any assessment themselves and there is generally very little trust between the psychologist and the (suspected) offender (Cima, 2003). Secondly, suspected offenders and forensic psychiatric patients can gain or lose quite a lot from the results of their assessment, such as the length of their prison sentence or whether they are ordered to undergo any kind of enforced treatment. Issues such as malingering (i.e., exaggerating negative qualities) or dissimulation and positive impression management (i.e., giving socially desirable answers) can certainly play a role in the way an individual presents him or herself (Cima, 2003; Wygant & Lareau, 2015) due to possible legal benefits. The presence of a (Cluster B) personality disorder or psychopathic personality characteristics also makes respondents more likely to yield biased assessments (Cima, 2003).

## Self-report instruments in forensic mental health assessment

Personality pathology can be assessed through a wide variety of assessment methods, including but not limited to self-report measures, observer-rated scales, and semi-structured interviews. Psychological assessment methods offer the forensic mental health practitioner thorough, qualitative and standardized results (Wygant & Lareau, 2015). Of the three above-mentioned methods, self-report assessments are the least complicated or time-consuming for forensic mental health professionals as they are quick and easy to administer and require most effort from the assessed, not the assessor.

Various studies have reported a relationship between personality traits pertaining to Cluster B personality disorders and deviant or disruptive behaviors using self-report assessment instruments (Boccaccini et al., 2010; Cunradi et al., 2009; Edens, Buffington-Vollum, Colwell, Johnson, & Johnson, 2002; Dolan & Blackburn, 2006; Hare, 2006; Neumann & Hare, 2008; Norlander & Eckhardt, 2005; Shechory et al., 2011; Taft et al., 2006; Walters, 2007). Miller and Lynam (2001) carried out a meta-analysis on the relationship between antisocial behaviors and basic dimensions of personality and found that the dimensions agreeableness and conscientiousness, from Costa and McCrae's Five-Factor Model (FFM; Costa & McCrae, 1990), showed the strongest association with antisocial behaviors. Further studies showed that individuals who commit crime or are psychopaths are generally low in agreeableness, exhibiting negative interpersonal and psychopathic characteristics such as deceitfulness, manipulateness, and a grandiose sense of self-worth, low in conscientiousness, meaning they lack responsibility and are unreliable (Miller & Lynam, 2001; Miller & Lynam, 2003; Miller, Lynam, Widiger, & Leukefeld 2001) and also display high levels of facets of neuroticism, pertaining to angry hostility and impulsiveness (Miller & Lynam, 2015; Widiger & Costa, 2012). The validity of self-report methods in forensic populations, however, has been questioned (Milton et al., 2005). Given the possible consequences for the forensic population of the outcome of their psychological evaluation, and the great diversity of self-report instruments available for personality traits – the majority of which were not designed specifically for forensic populations (Wygant & Lareau, 2015) – there appears to be little clarity or overview of the current level of knowledge on these maladaptive personality traits in forensic populations. Chapter Two discusses an investigation into the prevalence of negative, inflexible, and notable personality traits – as maladaptive or severe variants of the common dimensions of personality encompassed by the FFM – as assessed with self-report instruments, that lead to deviant or disruptive behavior and are most prevalent in forensic populations worldwide.

## Efficient assessment in forensic mental health evaluations

For various reasons it is important that the forensic mental health assessment process is efficient and cost-effective (Wygant & Lareau, 2015). In the field of general psychological practice, Widiger and Samuel (2005) recommend a two-step approach for an efficient assessment of personality psychopathology. This entails first administering a self-report questionnaire to screen for the potential presence of personality disorders, followed by a standardized (semi-)structured diagnostic interview to verify the presence of the disorder. This assessment process could pose many benefits for the forensic mental health field. If the screening questionnaire were quick and efficient, the amount of time required to confirm the presence or absence of a diagnosis would be shorter. However, the most important quality of self-report instruments in the assessment of psychopathology is that they are accurate and that the forensic evaluation report used to inform all relevant legal parties about the offender are clear, objective, and transparent (Wygant & Lareau, 2015).

Chapter Three discusses the suitability of the Dimensional Assessment of Personality Pathology – Short Form (DAPP-SF; Van Kampen, De Beurs, & Andrea, 2008) for this two-stage screening process in a psychiatric observation hospital where pre-trial psychological and psychiatric assessments are carried out on suspects of serious crimes. The DAPP-SF is a self-report instrument for the assessment of personality which has shown reliability and validity in the general population as well as in patients seeking treatment for personality disorders (Van Kampen et al., 2008) and in patients with mood, anxiety, and somatoform disorders (De Beurs, Rinne, Van Kampen, Verheul, & Andrea, 2009). The DAPP-SF proved to be able to

distinguish patients with personality disorders from the general population (De Beurs, Rinne, Van Kampen, Verheul, & Andrea, 2010). As this instrument is an accurate screening tool in the general population, it might also be so for forensic populations.

Continuing on the topic of self-report personality assessment instruments, the Minnesota Multiphasic Personality Inventory-2 (MMPI-2; Hathaway & McKinley, 1989; Nieberding et al., 2003) is one of the most widely used and researched self-report personality assessment instruments in correctional and forensic psychiatric settings. In a survey on the use of forensic assessment instruments, the majority of forensic psychological experts recommended the use of the MMPI-2 for evaluating an individual's mental state at the time of the offense (Lally, 2003), even though it has not been validated for forensic populations.

One popular line of research using the MMPI(-2) in the forensic arena has been the attempt to classify criminal offenders into distinct groups according to their profiles. Such classification may provide useful information about motives and for treatment and management for each specific subtype of offenders. Previous studies have found between two (Espelage et al., 2003; Hall, Graham, & Shepherd, 1991) and ten (Megargee, Carbonell, Bohn, & Sliger, 2001) different clusters. Studies indicating only two separate clusters raise doubt about the suitability of the MMPI-2 as the primary instrument for differentiating personality types in a forensic population.

Chapter Four discusses the replication of the above-mentioned previous studies aimed at deriving an empirical classification system using cluster analysis of MMPI-2 profiles of pretrial criminal defendants in a forensic psychiatric observation hospital, which could benefit legal decisions on level of criminal responsibility as well as treatment decisions.

## Personality disorder and criminal law

Once personality pathology has been established in pre-trial psychological assessments, judges have to decide on the level of criminal responsibility and corresponding level of punishment and/or treatment if the chance of recidivism is high. International jurisdictions differ on what disorders are deemed sufficient to warrant diminished criminal responsibility, criminal insanity and/or treatment of their mental health. In some jurisdictions, such as the U.S. and Canada, the presence of a personality disorder generally does not often lead to a verdict of insanity or diminished criminal responsibility (Rice & Harris, 1990; Warren, Murrie, Chauhan, Dietz & Morris, 2004). This could be partly due to the high prevalence of antisocial personality disorder in detained populations (Fazel & Danesh, 2002). A crucial diagnostic criterion for antisocial personality disorder is criminal versatility and repeated unlawful behaviors (American Psychiatric Association, 2000), making it almost by definition very prevalent in detained populations (Van der Wolf, 2012). It should be noted that under Dutch criminal law the crime the defendant has been accused of can never be used to form a diagnosis. In other words, the pre-trial diagnosis must also hold in the case of an acquittal. A history of criminal behavior in itself is not considered sufficient to warrant diminished criminal responsibility.

According to Sparr (2009), some argue that personality disorders are ever-present and rarely involve the cognitive defects required for an insanity plea or diminished criminal responsibility. Following this claim, one could argue that personality disorders do not greatly affect a person's freedom of will. Other authors have argued, however, that personality disorders and psychopathy *can* be interpreted as serious mental disorders, because they are based on developmental disabilities or deficits such as

cognitive deficiencies and biological impediments (Ciocchetti, 2003; Fine & Kennett, 2004; Herpertz & Sass, 2000; Mei-Tal, 2002; Palermo, 2007).

Given the complex role that personality disorder plays and the international variability regarding whether it should or shouldn't be a factor in determining the level of criminal responsibility by forensic mental health experts, Chapter Five discusses the Dutch forensic context. It presents two empirical investigations into whether personality disorders as well as psychopathic personality traits in criminal suspects are reasons for diminished criminal responsibility or enforced treatment in high security hospitals in the Netherlands. In a sample of suspects of serious crimes undergoing pre-trial assessment in a Dutch forensic psychiatric observation hospital, the first study examined the extent to which forensic mental health experts consider personality disorders in their recommendations on criminal responsibility and the need for enforced treatment. The second study investigated associations between scores on the Psychopathy Checklist-Revised (PCL-R; Hare, 2003) and expert recommendations on criminal responsibility and the need for enforced treatment in a similar sample at the same observation hospital.

## Personality disorder and enforced treatment

In the Netherlands, once the court has established that a mental disorder influenced an individual's free will and led to a crime, that individual's criminal responsibility can be considered diminished and the individual can be sentenced to enforced treatment in a high security hospital. As mentioned earlier, in Dutch forensic practice the presence of a personality disorder may lead to an offender being deemed diminished responsible and eligible for enforced treatment. The aim of enforced treatment is to protect society against individuals with a high risk of recidivism stemming from their disorder (De Kogel & Nagtegaal, 2006; Van Gemmert, Van Schijndel, Gordeau, & Casanova, 2013) and to facilitate the individual's gradual, safe, and justified return into society (Koenraadt, Mooij, & Van Mulbregt, 2007; Van der Wolf, 2012). As long as the forensic psychiatric patient's recidivism risk has not decreased sufficiently to warrant return to society, the treatment continues. This creates the possibility of lifelong detention in a high security setting. The median enforced treatment length for forensic psychiatric patients had gradually been increasing over the years, from 7.3 years for patients who entered the system in 1990 (Dienst Justitiële Inrichtingen, 2011) to its peak at 10.6 years for patients who entered the system in 1998 (Dienst Justitiële Inrichtingen, 2015). A recent study shows that the median enforced treatment length for patients who entered the system in 2006 has decreased somewhat to 8.0 years (Dienst Justitiële Inrichtingen, 2015).

Accompanied and unaccompanied leave are crucial steps in the treatment process. These periods of authorized freedom to move outside the secure zone of the forensic psychiatric hospital gradually prepare a patient for his or her return to society, by examining if treatment advances made in the hospital hold up in the less protective outside environment and if the patient can handle an increase in freedom and responsibility (Nagtegaal et al., 2011; Bernstein, Nijman, Karos, Keulen-De Vos, De Vogel, & Lucker, 2012). After a number of serious incidents with forensic patients in the past, forensic psychiatric hospitals have become hesitant in granting leave, and are prolonging first applications for leave of their patients (Mevis, 2011). At the same time, there is increasing pressure from the government to speed up the leave process, and with that, shorten the total treatment length. Forensic hospitals are therefore constantly facing deliberations and complex decisions (Nijman,

De Kruyk, & Van Nieuwenhuizen, 2002) concerning suitability of leave, conditional release and the potentially harmful consequences of a new offense (Cima, 2003). Expanding on knowledge about personality disorder as a prognostic factor for treatment can contribute to more effective forensic treatment.

According to De Kogel and Nagtegaal (2006), one of the important predictive factors for the success of forensic treatment of a disorder concerns the characteristics of the person being treated. Research on the relationship between personality pathology and treatment length for forensic patients, however, is limited. As far as the authors are aware, longitudinal or prospective research on treatment length for patients with personality disorder in enforced forensic treatment has never been carried out. Chapter Six discusses an investigation into the predictive value of personality disorder on treatment length, at first leave request and end of treatment, of offenders detained in high security forensic treatment hospitals in the Netherlands. This study combined data from pre-trial assessment – from a population with 76.5% personality disorder prevalence – with data containing treatment characteristics from forensic hospitals. It aimed to elaborate on current knowledge regarding the influence of personality disorder on treatment length of forensic patients in the Netherlands. To examine possible confounding, the effect of index offense, Axis I disorder, substance abuse history and intellectual functioning were also investigated.

## Research questions

In summary, the present dissertation investigates the following five issues concerning the prevalence, assessment and treatment of personality pathology – i.e., both personality disorders and maladaptive personality traits – in forensic practice:

- 1) Given the possible consequences for the forensic population of the outcome of their psychological evaluation, and the fact that the majority of self-report instruments available for personality traits were not designed specifically for forensic populations, there is a need for clarity on the prevalence of these maladaptive personality traits in forensic populations. What are the self-reported levels of antisocial and psychopathic personality traits as well as traits associated with agreeableness, conscientiousness and neuroticism in forensic populations, i.e. anger, aggression, hostility, and impulsivity in forensic populations, compared to normal populations? (Chapter Two)
- 2) When applying a two-step approach for an efficient assessment of personality psychopathology in a forensic mental health setting, can the DAPP-SF be used as a screening tool for personality disorder and can it correctly determine who should and should not undergo a standardized semi-structured diagnostic interview to verify the presence of the disorder, using data from a forensic psychiatric hospital where pre-trial evaluations are carried out on suspects of serious crimes? (Chapter Three)

- 3) When assessing personality traits in a known heterogeneous population of pretrial criminal defendants of serious crimes in a forensic psychiatric observation hospital, can a cluster analysis of MMPI-2 profiles produce a number of distinct personality profiles? (Chapter Four)
- 4) Once a personality disorder has been established in pre-trial psychiatric forensic evaluations, how do Dutch mental health experts consider this diagnosis in their recommendations regarding criminal responsibility compared to other psychiatric conditions, and how do they advise on the necessity of enforced treatment in a high security hospital? Furthermore, how do they consider various aspects of psychopathy as measured by the Psychopathy Checklist-Revised (PCL-R) into these judgments? (Chapter Five)
- 5) For those offenders who have been committed to enforced treatment in a high security hospital, what is the predictive value of a personality disorder for treatment duration at first leave request and end of treatment? Are there any confounding effects of index offense, Axis 1 disorder, substance abuse history and intellectual functioning? (Chapter Six)

## Research method

The current dissertation investigates the prevalence, assessment and prognostic value for length of treatment of personality pathology, or both personality disorders and maladaptive personality traits, in a forensic mental health setting. While Chapter Two presents meta-analytic data based on literature review, the research presented in Chapters Three, Four, and Five was all carried out using the database of results on assessment instruments acquired after the revision of the assessment process in the PBC. The research presented in Chapter Six is one of the first of its kind to merge data from digital databases from different sources available in the Dutch judicial system and over a range of treatment hospitals.





Chapter Two  
*Self-reported  
personality traits in  
forensic populations:  
A meta-analysis*

## Abstract

The current study covers a systematic review and meta-analysis of the prevalence of self-reported deviant or disruptive personality traits: anger, aggression, hostility, antisocial traits, psychopathy, and impulsivity in forensic populations worldwide. A computer-based search of titles was carried out using the PubMed electronic database for articles published in English that included a self-report instrument for personality characteristics in combination with a forensic population (i.e., detained in remand, sentenced and/or in enforced treatment, or on parole). The final sample consisted of 39 studies (N = 11,716) that together used 17 different instruments and reported on 32 subscales or constructs that fitted our current interest. Results showed significantly higher levels of self-reported antisocial and psychopathic features in forensic samples, including a significant effect of the assessment instrument and subscale used. No significant differences were found for self-reported impulsivity, anger, aggression, or hostility in forensic populations compared to norm scores of non-forensic samples. Possible explanations, including suggestions that forensic populations are prone to providing socially desirable answers on self-report questionnaires, possibly to gain advantages such as a lower prison sentence or to avoid enforced treatment, are discussed, as well as limitations, and suggestions for future research and clinical practice.<sup>1</sup>

## Introduction

According to a systematic review carried out by Fazel and Danesh in 2002, 65% of the male general prison population and 42% of the female general prison population in western countries has a personality disorder, based on interview methods of assessing these disorders. The authors also found that prisoners are ten times more likely to have antisocial personality disorder than the general population, with a prevalence of 47% in men and of 21% in women.

Many of the separate personality traits that correspond to antisocial personality disorder and narcissistic personality disorder have been studied individually and shown to have a relationship with deviant or disruptive behaviors – such as antisocial features (Edens et al., 2002; Shechory et al., 2013), psychopathic features (Dolan & Blackburn, 2006; Edens et al., 2002; Hare, 2006; Neumann & Hare, 2008), anger (Norlander & Eckhardt, 2005; Taft et al., 2006), aggression (Boccaccini et al., 2010; Dolan & Blackburn, 2006; Shechory et al., 2013; Walters, 2007), hostility (Dolan & Blackburn, 2006; Norlander & Eckhardt, 2005), impulsivity (Cunradi et al., 2009), and dominance (Dolan & Blackburn, 2006).

In order to contribute to understanding the role of personality in crime within the field of criminology, Miller and Lynam (2001) carried out a meta-analysis on the relationship between a broad interpretation of antisocial behaviors and basic dimensions of personality. Results of 59 studies indicated that the dimensions agreeableness and conscientiousness, from Costa and McCrae's Five-Factor Model (FFM; Costa & McCrae, 1990), showed the strongest association with antisocial behaviors. Similarly, Miller, Lynam, Widiger, and Leukefeld (2001) studied the relationship between psychopathy and the FFM. Studies show that individuals who commit crime or are psychopaths are generally low in agreeableness, exhibiting negative interpersonal and psychopathic characteristics such as deceitfulness, manipulateness, and a grandiose sense of self-worth, and low in conscientiousness – meaning they lack responsibility and are unreliable (Miller & Lynam, 2001; Miller & Lynam, 2003; Miller et al., 2001) and also display high levels of facets of neuroticism, pertaining to angry hostility and impulsiveness (Miller & Lynam, 2015; Widiger & Costa, 2012).

In line with these above-mentioned studies, the current study examines negative, inflexible, and

.....  
<sup>1</sup> Spaans, M., Molendijk, M. L., De Beurs, E., Rinne, T., & Spinhoven, P. (2016). Self-reported personality traits in forensic populations: A meta-analysis. *Psychology, Crime & Law*. Advance online publication.

notable personality traits – as maladaptive or severe variants of the common dimensions of personality encompassed by the FFM – that together have the potential to cause antisocial or criminal behavior and the accompanying likelihood of considerable damage or distress to persons and society. The importance of studying separate personality traits is stressed in earlier versions of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR; American Psychiatric Association, 2000) which provided for the opportunity to record maladaptive personality traits that may be below the threshold of a disorder, but still of great diagnostic importance. Examining severe variants of common personality dimensions, even when they may not constitute a personality disorder, is also more in line with the proposed dimensional approach to personality disorder as presented in the DSM-5 (American Psychiatric Association, 2013a). The aim of this study was to contribute to finding the best way to assess, describe, and diagnose severe variants of common personality traits and the ensuing antisocial or criminal behavior – in order to add to existing knowledge on the role of personality in crime, to be able to better treat personality disordered individuals, and ideally to eventually prevent crime.

### **Self-report method**

In order to further study the severe variants of common personality traits that accompany antisocial behavior in forensic populations – such as those belonging to DSM-5 Cluster B antisocial personality disorder and narcissistic personality disorder – the actual extent to which they are present should first be assessed. Ideally this is also done within forensically relevant subgroups differing in age, gender, and type of crime committed. The above-mentioned selection of studies used a wide variety of assessment methods to evaluate maladaptive or severe personality traits, ranging from self-report measures to observer-rated scales and semi-structured interviews. Of these methods, self-report assessments are the least complicated and least time-consuming. However, the validity of self-report methods in forensic populations that show the most deviant or disruptive behaviors is questionable, partly due to the fact that forensic psychiatric patients can have a lot to gain or lose from the results of their assessment, such as the length of their prison sentence or whether or not they receive enforced treatment (Milton et al., 2005; Spaans, Barendregt, Muller, De Beurs, Nijman, & Rinne, 2009; Spaans, Rinne, De Beurs, & Spinhoven, 2015). Moreover, given the great diversity of self-report instruments for personality traits, there appears to be little clarity or overview of the current level of knowledge on the subject. The aim of the present study is therefore to review which personality traits, assessed with self-report measures, are most prevalent in forensic settings.

The current study investigates antisocial behavior and psychopathic features and the severe variants of common personality traits of the two most relevant personality dimensions of agreeableness, neuroticism, and conscientiousness in forensic populations, i.e. self-reported anger, aggression, hostility, and respectively, impulsivity.

## **Method**

A computer-based search of titles was carried out using the PubMed electronic database for articles published in English that included a self-report instrument for personality characteristics in combination with a forensic population (i.e., detained in remand, sentenced and/or in enforced treatment, or on parole). The search strategy for the current study is presented in Appendix 1. There were no geographical and/or cultural restrictions, or restrictions on time period in which the studies

were conducted or research design. This search strategy resulted in 2,840 potential articles, published between 1946 and 2015. Further possibly relevant publications were obtained from reference lists. Titles and abstracts were screened for appropriateness regarding inclusion in the current study by MS and a research assistant. Exclusion of articles was discussed between the two and in cases of disagreement MLM was consulted. Along with 29 additional articles identified through other sources, a total of 187 articles were selected for full text assessment in more detail. These articles contained a total of 180 different subscales to measure personality characteristics with. Studies were excluded if they were written in a language other than English, if the study population was juvenile or adolescent, if the study did not assess a personality trait related to anger, aggression, hostility, antisocial traits, psychopathy, and impulsivity, if the study population was not suspected of or charged with a crime, or if the (self-report) assessment instrument used in the study was not specified or validated. As anger was considered to be a personality trait within the context of the current study (Martin, Watson, & Wan, 2000), assessment instruments were chosen that contained subscales pertaining to trait anger.

### **Extraction of data**

Initially, 58 articles were selected for the study, based on whether they reported mean scores on (sub) scales corresponding to anger, aggression, hostility, antisocial traits, psychopathy, and impulsivity in forensic populations. Authors of articles that did not cite sample size, mean, and/or standard deviations of personality characteristics were contacted via e-mail and were asked to provide this data. Seven studies for which the relevant data could not be collected this way were excluded from further analyses.

A small number of studies ( $n = 5$ ) reported mean personality scores for a forensic population as well as for healthy controls from the same study. In those cases, the norm or reference scores were extracted from the article. In the other 34 cases where there was no mention of controls generated from the general or healthy population in the same study, manuals were requested from the corresponding publishers in order to compare mean scores of forensic population to norm scores of the self-report instruments. For those that could not be provided, literature searches were carried out to find published articles containing the original norm or reference scores for the particular self-report instruments. Where no original norm or reference scores could be found, other available publications were used for reference scores. The highest available match for country of origin, versions of the self-report instruments, year of study, and socio-demographic variables of the forensic sample was chosen where possible. The most preferable sources for this information were introductions to the instrument and validation studies. Instruments for which neither a manual nor relevant publications could be found were excluded from this study, as well as one case in which the norm or reference scores that were found through the above-mentioned method were presented in very wide ranges of normalized T scores that could not be matched to the exact T scores given in the study and no further information could be found or provided. The source of the norm or reference scores was included in the meta-analysis as a potential moderator variable: derived from the original article, general norms derived from manuals, or reference scores derived from other publications.

In order to obtain sufficient sample size per personality trait for analysis, it was decided to group the personality traits together into three central categories: (1) Antisocial/Psychopathy, (2) Anger/Aggression/Hostility, and (3) Impulsivity. This corresponds closely with the conceptualizing of severe variants of personality traits by Miller and Lynam (2001; 2003; 2015), Miller and colleagues (2001), and Widiger and Costa (2012), with the first category representing features of antisocial and psychopathic

behavior in general and the latter two categories representing the more negative features of the FFM dimensions agreeableness and conscientiousness, respectively.

Means, standard deviations, sample sizes, and data concerning potential moderators were extracted from each article by MS and a research assistant. Categorical variables included the name or kind of the assessment instrument, the name or kind of subscale or construct used, the judicial phase the study sample was in (i.e., remanded, convicted, probationed, and/or in treatment), gender distribution of the sample (only males, only females, or mixed), country of origin, source of the reference or control means, standard deviations, sample sizes, the type of crime(s) for which the study sample was incarcerated, and the numerical variables mean age in years and year of study.

The final sample consisted of 39 studies that together used 17 different instruments and 32 subscales or constructs. The search process is shown in a flow diagram in Appendix 2. Table 1 shows the three categories of personality traits along with corresponding constructs or subscales in alphabetical order, assessment instrument, and citation(s) per construct.

### **Analysis**

Comprehensive Meta-Analysis (CMA; Borenstein, Hedges, Higgins, & Rothstein, 2005) software was used to calculate pooled effect size estimates over studies. Moderator analyses and meta-regression analyses were carried out with the above-mentioned potential moderator variables. With regard to the moderator analyses, outcomes from multiple subgroups within the same study were treated as not completely independent. Methodological quality of the included studies was assessed using an adapted version of the Newcastle-Ottawa Scale for assessing the quality of non-randomized studies in meta-analyses (NOS; Wells et al., 2015). Publication bias was examined for each outcome category by means of funnel-plots.

**Table 1.** Personality trait categories and corresponding instruments and subscales

Category	Subscale	Instrument	Study name(s)
Anger/ Aggression/ Hostility	Aggression	EPPS (Edwards 1959)	1. Reith et al. 1975
	Aggression	PAI (Morey 1991)	1. Boccaccini et al. 2010 2. Haden & Shiva 2008 3. Laulik et al. 2007 4. Magaletta et al. 2014 5. Magyar et al. 2012 6. Newberry & Shuker 2012 7. Percosky et al. 2013 8. Ruiz et al. 2014
	Aggression	SNAP (Clark 1996)	1. Hurt & Oltmanns 2002
	Anger	AQ (Buss & Perry 1992)	1. Hulme & Middleton 2013
	Anger in	AX (Spielberger Johnson, Jacobs Krasner Oesterle & Worden 1986)	1. Dear et al. 2003 2. Kalichman 1990 3. Kroner & Reddon 1992
	Anger out	AX (Spielberger et al. 1986)	1. Dear et al. 2003 2. Kalichman 1990 3. Kroner & Reddon 1992
	Anger control	AX (Spielberger et al. 1986)	1. Dear et al. 2003 2. Kalichman 1990 3. Kroner & Reddon 1992
	Anger control	STAXI (Spielberger 1988)	1. Blagov et al. 2011 2. Fernández-Montalvo et al. 2012 3. Roy et al. 2014
	Anger in	STAXI (Spielberger 1988)	1. Blagov et al. 2011 2. Fernández-Montalvo et al. 2012 3. Roy et al. 2014
	Anger out	STAXI (Spielberger 1988)	1. Blagov et al. 2011 2. Fernández-Montalvo et al. 2012 3. Roy et al. 2014
	Hostility	BDHI (Buss & Durkee 1957)	1. Firestone et al. 1998 2. Knust & Stewart 2002 3. Roy et al. 2014
	Hostility	AQ (Buss & Perry 1992)	1. Hulme & Middleton 2013
	Physical aggression	AQ (Buss & Perry 1992)	1. Hulme & Middleton 2013

Category	Subscale	Instrument	Study name(s)
	Trait anger	STAS (Spielberger Jacobs Russel & Crane, 1983)	1. Kalichman 1990 2. Kroner & Reddon 1992
	Trait anger	STAXI (Spielberger 1988)	1. Echeburúa et al. 2003 2. Fernández-Montalvo et al. 2012 3. Shorey et al. 2011
	Trait anger reaction	STAXI (Spielberger 1988)	1. Fernández-Montalvo et al. 2012
	Trait anger temperament	STAXI (Spielberger 1988)	1. Fernández-Montalvo et al. 2012
	Verbal aggression	AQ (Buss & Perry 1992)	1. Hulme & Middleton 2013
Antisocial/ Psychopathy	Antisocial features	PAI (Morey 1991)	1. Boccaccini et al. 2010 2. Edens 2009 3. Haden & Shiva 2008 4. Laulik et al. 2007 5. Magaletta et al. 2014 6. Magyar et al. 2012 7. Newberry & Shuker 2012 8. Percosky et al. 2013 9. Ruiz et al. 2014
	Dominance	EPPS (Edwards 1959)	1. Reith et al. 1975
	Dominance	PAI (Morey 1991)	1. Boccaccini et al. 2010 2. Edens 2009 3. Laulik et al. 2007 4. Magaletta et al. 2014 5. Magyar et al. 2012 6. Newberry & Shuker 2012 7. Percosky et al. 2013 8. Ruiz et al. 2014
	Empathic concern	IRI (Davis 1980)	1. Hepper et al. 2014
	Empathy	BEES (Mehrabian & Epstein 1972)	1. Hulme & Middleton 2013
	Manipulativeness	SNAP (Clark 1996)	1. Hurt & Oltmanns 2002

Category	Subscale	Instrument	Study name(s)
	Psychopathic deviate	MMPI-1 (Hathaway & McKinley 1942)	1. Adams 1976 2. Bauer & Clark 1976 3. Holland & Holt 1975 4. Kalichman 1990 5. McCreary & Padilla 1977 6. Panton 1976 7. Roman & Gerbing 1989 8. Scott & Conn 1979 9. Twomey & Hendry 1969 10. Walls et al. 1977
	Psychopathic deviate	MMPI-168 Overall & Gomez-Mont 1974)	1. Scott & Conn 1979 2. Valliant et al. 2000 3. Valliant et al. 2004 4. Walls et al. 1977
	Psychopathic deviate	MMPI-2 (Hathaway & McKinley 1989)	1. Claes et al. 2014 2. Shechory et al. 2013
Impulsivity	Impulsivity	EIS (Eysenck Pearson Easting & Allsopp 1985)	1. Shorey et al. 2011
	Impulsivity	BIS-10 (Barratt 1985)	1. Echeburúa et al. 2003 2. Herpertz et al. 2001
	Impulsivity	BIS-11 (Patton Stanford Barratt 1995)	1. Cuomo et al. 2008 2. Haden & Shiva 2008 3. Illiceto et al. 2012
	Impulsivity	EASI Buss & Plomin 1975	1. Blagov et al. 2011
	Impulsivity	SNAP (Clark 1996)	1. Hurt & Oltmanns 2002

## Results

Table 2 provides the 39 included studies (total  $N = 11,716$ ) along with instrument(s) and subscale(s) used and sample size by study. Results of the meta-analysis are presented for each overall category separately.

**Table 2.** *Articles used in meta-analysis*

Study name	Instrument(s) & subscale(s)	Sample size (n)
Adams 1976	1. MMPI-1 (psychopathic deviate)	100
Bauer & Clark 1976	1. MMPI-1 (psychopathic deviate)	88
Blagov et al. 2011	1. EASI (impulsivity) 2. STAXI (anger expression in) 3. STAXI (anger expression out) 4. STAXI (anger expression control)	33 34 34 34
Boccaccini et al. 2010	1. PAI (antisocial) 2. PAI (aggression) 3. PAI (dominance)	1412 1412 1412
Claes et al. 2014	1. MMPI-2 (psychopathic deviate)	110
Cuomo et al. 2008	1. BIS-11 (total score)	903
Dear et al. 2003	1. AX (anger in) 2. AX (anger out) 3. AX (anger control)	397 397 397
Echeburúa et al. 2003	1. STAXI (trait anger) 2. BIS-10 (total score)	54 54
Edens 2009	1. PAI (antisocial) 2. PAI (dominance)	1062 1062
10. Fernández-Montalvo et al. 2012	1. STAXI (trait anger) 2. STAXI (anger in) 3. STAXI (anger out) 4. STAXI (trait anger temperament) 5. STAXI (trait anger reaction)	123 123 123 123 123
11. Firestone et al. 1998	1. BDHI (total score)	96
12. Haden & Shiva 2008	1. BIS-11 (total score) 2. PAI (antisocial) 3. PAI (aggression)	436 139 139
13. Hepper et al. 2014	1. IRI (empathic concern)	77
14. Herpertz et al. 2001	1. BIS-10 (total score)	43
15. Holland & Holt 1975	1. MMPI-1 (psychopathic deviate)	295
16. Hulme & Middleton 2013	1. AQ (anger) 2. AQ (hostility) 3. AQ (physical aggression) 4. AQ (verbal aggression) 5. BEES (empathy)	38 38 38 38 38

Study name	Instrument(s) & subscale(s)	Sample size (n)
17. Hurt & Oltmanns, 2002	1. SNAP (aggression) 2. SNAP (manipulativeness) 3. SNAP (impulsivity)	157 157 157
18. Iliceto et al. 2012	1. BIS-11 (total score)	40
19. Kalichman 1990	1. AX (anger in) 2. AX (anger out) 3. AX (anger control) 4. MMPI-1 (psychopathic deviate) 5. STAS (trait-anger)	111 111 111 111 111
20. Knust & Stewart 2002	1. BDHI (total score)	92
21. Kroner & Reddon 1992	1. AX (anger in) 2. AX (anger out) 3. AX (anger control)	137 137 137
22. Laulik et al. 2007	1. PAI (antisocial) 2. PAI (aggression) 3. PAI (dominance)	30 30 30
23. Magaletta et al. 2014	1. PAI (antisocial) 2. PAI (aggression) 3. PAI (dominance)	61 61 61
24. Magyar et al. 2012	1. PAI (antisocial) 2. PAI (aggression) 3. PAI (dominance)	331 331 331
25. McCreary & Padilla 1977	1. MMPI-1 (psychopathic deviate)	304
26. Newberry & Shuker 2012	1. PAI (antisocial) 2. PAI (aggression) 3. PAI (dominance)	268 268 268
27. Panton 1976	1. MMPI-1 (psychopathic deviate)	2585
28. Percosky et al. 2013	1. PAI (antisocial) 2. PAI (aggression) 3. PAI (dominance)	34 34 34
29. Reith et al. 1975	1. EPPS (dominance) 2. EPPS (aggression)	140 140
30. Roman & Gerbing 1989	1. MMPI-1 (psychopathic deviate)	340
31. Roy et al. 2014	1. BDHI (total score) 2. STAXI (anger in) 3. STAXI (anger out) 4. STAXI (anger control) 5. STAXI (trait anger)	1537 1537 1537 1537 1537
32. Ruiz et al. 2014	1. PAI (antisocial) 2. PAI (aggression) 3. PAI (dominance)	124 124 124
33. Scott & Conn 1979	1. MMPI-1 (psychopathic deviate) 2. MMPI-168 (psychopathic deviate)	165 165
34. Shechory et al. 2013	1. MMPI-2 (psychopathic deviate)	230

Study name	Instrument(s) & subscale(s)	Sample size (n)
35. Shorey et al. 2011	1. EIS (impulsivity) 2. STAXI (trait anger)	80 80
36. Twomey & Hendry 1969	1. MMPI-1 (psychopathic deviate)	214
37. Valliant et al. 2000	1. MMPI-168 (psychopathic deviate)	54
38. Valliant et al. 2004	1. MMPI-168 (psychopathic deviate)	88
39. Walls et al. 1977	1. MMPI-1 (psychopathic deviate) 2. MMPI-168 (psychopathic deviate)	98 98

Note. AQ = Aggression Questionnaire. AX = Anger Expression. BDHI = Buss Durkee Hostility Inventory. BEES = Balanced Emotional Empathy Test. BIS = Barratt Impulsiveness Scale. EASI = Emotionality, Activity, Sociability, and Impulsivity. EIS = Eysenck Impulsiveness Scale. IRI = Interpersonal Reactivity Index. MMPI = Minnesota Multiphasic Personality Inventory. PAI = Personality Assessment Inventory. SNAP = Schedule for Nonadaptive and Adaptive Personality. STAS = State Anger Scale. STAXI = State-Trait Anger Expression Inventory.

### Antisocial/Psychopathy

Random effect models for the 27 studies ( $n = 8,263$ ) that assessed self-reported antisocial and psychopathic features showed higher scores on these traits in forensic samples compared to reference populations with a large effect size ( $g = 1.05$ , 95% CI [0.78, 1.32],  $p < .001$ ). Results also showed between-study heterogeneity ( $I^2 = 98.87$ ,  $Q = 5130.35$ ,  $p < .001$ ).

The studies used the Minnesota Multiphasic Personality Inventory (MMPI) versions 1 and/or 168, the MMPI-2, the Personality Assessment Inventory (PAI), the Balanced Emotional Empathy Test (BEES), Edwards Personal Preference Schedule (EPPS), Interpersonal Reactivity Index (IRI), and Schedule for Nonadaptive and Adaptive Personality (SNAP). Moderator analyses showed a significant moderation effect of the assessment instrument ( $Q = 94.34$ ,  $df = 7$ ,  $p < .001$ ;  $g = 0.07$ , 95% CI [-0.03, 0.16]) and subscale used ( $Q = 88.45$ ,  $df = 5$ ,  $p < .001$ ;  $g = 0.05$ , 95% CI [-0.05, 0.15]; see Table 3 for effect sizes), indicating slightly different self-reported levels of antisocial and psychopathic features depending on which instrument or subscale was used. Due to the significant and large effects of all versions of the MMPI, and given the suggestion of restricted usefulness of the MMPI(-2) within a forensic context by Spaans and colleagues (2009), it was decided to investigate a possible difference in self-reported levels of antisocial and psychopathic features between the MMPI and other instruments. Results showed that with the division of the instruments used into MMPI vs. other instruments, the significant effect of the assessment instrument used remained ( $Q = 12.71$ ,  $df = 1$ ,  $p < .001$ ), with a large effect size of 1.39 ( $p < .001$ ) for all versions of the MMPI, compared to a moderate effect size of 0.46 ( $p < .001$ ) for all other instruments. This indicated different levels of self-reported antisocial and psychopathic traits when using the MMPI compared to other instruments.

Moderator analyses of country of sample origin showed a significant moderation effect ( $Q = 129.40$ ,  $df = 4$ ,  $p < .001$ ;  $g = 0.52$ , 95% CI [0.43, 0.62]). Effect sizes and significance per country of sample origin, as shown in Table 3, indicated that levels of self-reported antisocial and psychopathic traits differ between forensic populations and reference populations in Belgium, the United States, and Israel. Gender showed a large significant overall effect ( $Q = 76.63$ ,  $df = 3$ ,  $p < .001$ ;  $g = 1.53$ , 95% CI [1.39, 1.66]), indicating significantly different levels of self-reported antisocial and psychopathic traits between the male and mixed samples in forensic populations and reference populations. Given that the large pooled effect size for female samples ( $g = 1.41$ ) was based on only two studies, this is most likely due to low power (see Table 3). There were several categories for judicial phase within the samples (see

Table 3). Judicial phase showed a significant moderation effect ( $Q = 43.06$ ,  $df = 5$ ,  $p < .001$ ;  $g = 0.57$ , 95% CI [0.51, 0.63]). Results indicated that for every judicial phase, significant differences in level of psychopathic or antisocial traits were observed between the forensic samples and the reference populations (see Table 3).

A moderator analysis for type of crime revealed a large significant effect on levels of self-reported antisocial and psychopathic traits ( $Q = 96.81$ ,  $df = 6$ ,  $p < .001$ ;  $g = 1.55$ , 95% CI [1.45, 1.66]), indicating differences between forensic populations that were incarcerated for felonies, misdemeanors, sex offenses, and other serious offenses compared to non-forensic populations. Eight studies did not specify the type of crime for which their study sample was incarcerated. Results for categories of source of the control or reference scores showed a large significant moderation effect ( $Q = 8.58$ ,  $df = 1$ ,  $p < .001$ ;  $g = 0.91$ , 95% CI [0.67, 1.15]), indicating differences in levels of self-reported antisocial and psychopathic traits when means and standard deviations were compared to previously published general norms versus comparisons with a control group in the corresponding study (see Table 3).

Meta-regression analysis on mean age of the forensic sample and year of study showed no significant effect of these possible moderator variables (Coeff. =  $-.02$ ,  $p = .46$ ; Coeff. =  $-.00$ ,  $p = .67$ , respectively). Inspection of the funnel plot for publication bias showed an even distribution of the 27 studies across the combined effect size, indicating that a meaningful publication bias was unlikely.

**Table 3.** *Antisocial/Psychopathy: Results of moderator analyses (27 studies)*

Instrument name	No. of studies	Hedge's g	95% CI LL	95% CI UL
BEES	1	0.14	-0.23	0.51
EPPS	1	-0.13	-0.29	0.04
IRI	1	-0.50**	-0.82	-0.17
MMPI-1	10	1.68**	1.12	2.23
MMPI-168	4	0.92**	0.29	1.55
MMPI-2	2	0.95**	0.35	1.54
PAI	9	1.05**	0.68	1.43
SNAP	1	-0.09	0.26	0.08
Subscale name	No. of studies	Hedge's g	95% CI LL	95% CI UL
Antisocial features	9	1.05**	0.68	1.43
Dominance	9	-0.13	-0.29	0.04
Empathic concern	1	-0.50**	-0.82	-0.17
Empathy	1	0.14	-0.23	0.51
Manipulativeness	1	-0.09	-0.26	0.08
Psychopathic Deviate	14	1.38**	0.98	1.78
Sample origin	No. of studies	Hedge's g	95% CI LL	95% CI UL
Belgium	1	2.14**	1.79	2.49
Canada	3	0.14	-0.08	0.35
Israel	1	0.36**	0.25	0.48
United Kingdom	3	0.44	-0.49	1.37

Instrument name	No. of studies	Hedge's <i>g</i>	95% CI LL	95% CI UL
United States	19	1.50**	1.13	1.87
<i>Gender</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Female	2	1.42	-0.19	3.02
Male	23	0.96**	0.68	1.25
Male and female	1	1.30**	1.11	1.49
<i>Judicial phase</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Convicted	17	1.11**	0.74	1.48
Convicted and in treatment	3	1.30**	0.40	2.21
Probationed and in treatment	2	0.35**	0.24	0.46
Remanded and/or convicted	1	0.61**	0.54	0.69
Remanded and/or convicted and in treatment	3	1.61**	1.06	2.17
<i>Type of crime</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Felonies	3	1.61**	0.95	2.27
Internet sex offenses	2	0.84	-0.34	1.71
Misdemeanors	1	1.87**	1.74	2.00
Mixed	8	0.58**	0.34	0.81
Serious offenses	2	1.88**	1.39	2.36
Sex offenses	4	1.66**	0.81	2.50
<i>Source of reference scores</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
General norms	24	1.15**	0.86	1.43
Original article	3	0.36	-0.09	0.80

Note. CI = Confidence interval. LL = Lower limit. UL = Upper limit.

\* =  $p < .05$ . \*\* =  $p < .01$

### Anger/Aggression/Hostility

No statistically significant overall effect was found for self-reported anger, aggression and hostility in forensic populations ( $g = .19$ , 95% CI [-.03, .42],  $p = .09$ ) for the 21 corresponding studies ( $n = 5,692$ ). Between-study heterogeneity was found ( $I^2 = 97.07$ ,  $Q = 1329.719$ ,  $p < .001$ ).

The studies used the Personality Assessment Inventory (PAI), the Edwards Personal Preference Schedule (EPPS), the Anger Expression (AX) Scale, the State-Trait Anger Expression Inventory (STAXI), the State-Trait Anger Scale (STAS), the Buss Durkee Hostility Inventory (BDHI), the Aggression Questionnaire (AQ), and the Schedule for Nonadaptive and Adaptive Personality (SNAP). A moderator analysis showed a significant effect of the assessment instrument used ( $Q = 38.92$ ,  $df = 7$ ,  $p < .001$ ), although the point estimate corresponding to the overall effect was 0.09, indicating a very small effect. Table 4 shows the results per instrument used, including effects for the AX, EPPS, SNAP, and STAXI. A moderator analysis for subscale used showed no significant effect ( $Q = 14.80$ ,  $df = 10$ ,  $p < .14$ ; see Table 4 for effect sizes). More than half of the subscales used were named 'aggression' (52.9%). The point estimate corresponding to the overall effect was 0.01, showing almost no effect.

Moderator analyses of country of sample origin showed a small significant overall moderation effect ( $Q = 53.41$ ,  $df = 5$ ,  $p < .001$ ;  $g = -.13$ , 95% CI [-0.22, 0.04]). Effect sizes per country are shown in Table 4 and indicate differences in levels of self-reported anger, aggression, and hostility in Spanish, Italian, and Canadian forensic populations compared to reference populations. Gender also showed a small significant moderation effect ( $Q = 94.50$ ,  $df = 2$ ,  $p < .001$ ;  $g = 0.16$ , 95% CI [0.05, 0.26]). Effect sizes per gender group indicate significant differences in levels of self-reported anger, aggression, and hostility in female forensic populations and in mixed samples (see Table 4). Judicial phase showed a moderate significant effect ( $Q = 126.66$ ,  $df = 5$ ,  $p < .001$ ;  $g = 0.44$ , 95% CI [0.34, 0.53]). Effect sizes and significance per judicial phase indicate slightly lower levels of self-reported anger, aggression, and hostility in forensic populations on probation and in treatment at the same time, in populations in remand, and in samples with a mix of remanded and convicted offenders receiving treatment, compared to non-imprisoned populations (see Table 4).

Six studies did not provide information the specific type of crime. Overall results of the moderator analysis showed a moderately negative significant effect ( $Q = 112.98$ ,  $df = 5$ ,  $p < .001$ ;  $g = -0.68$ , 95% CI [-0.83, 0.53]). Different levels of anger, aggression, and hostility were self-reported in forensic populations that were incarcerated for gender violence. Only one study featured reference scores of a non-incarcerated sample alongside forensic scores. Results of the moderator analysis showed a small significant effect ( $Q = 16.39$ ,  $df = 1$ ,  $p < .001$ ;  $g = 0.35$ , 95% CI [0.16, 0.54]). Effect sizes for the two sources of the reference scores are shown in Table 4.

Meta-regression analysis on mean age of the forensic sample and year of study showed no significant effects of these possible moderator variables (Coeff. =  $-.02$ ,  $p = .38$ ; Coeff. =  $-.01$ ,  $p = .10$ , respectively). Inspection of the funnel plot for publication bias showed an even distribution of the 21 studies across the combined effect size, again indicating that a meaningful publication bias was unlikely.

**Table 4.** Anger/Aggression/Hostility: Results of moderator analyses (21 studies)

Instrument name	No. of studies	Hedge's g	95% CI LL	95% CI UL
AQ	1	0.09	-0.24	0.42
AX	3	0.33*	0.04	0.62
BDHI	3	0.15	-0.60	0.90
EPPS	1	0.35**	0.13	0.58
PAI	8	0.32	-0.16	0.79
SNAP	1	-0.22*	-0.40	-0.04
STAXI	5	-1.20*	-2.05	-0.36
Subscale name	No. of studies	Hedge's g	95% CI LL	95% CI UL
Aggression	10	0.30	-0.01	0.62
Anger	1	0.09	-0.24	0.42
Anger control	5	-0.13	-0.63	0.36
Anger out	6	-1.54	-1.76	-1.31
Hostility	4	0.15	-0.60	0.90
Trait anger	7	-0.64	-1.33	0.05

Instrument name	No. of studies	Hedge's g	95% CI LL	95% CI UL
<i>Sample origin</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Australia	2	0.42	-0.40	1.23
Canada	3	0.22*	0.02	0.43
Italy	1	-0.13*	-0.24	-0.02
Spain	1	-1.28**	-1.80	-0.75
United Kingdom	2	0.14	-1.08	1.37
United States	11	0.05	-0.36	0.46
<i>Gender</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Female	2	-0.25**	-0.39	-0.10
Male	18	0.05	-0.21	0.31
Male and female	1	0.91**	0.72	1.09
<i>Judicial phase</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Convicted	10	-0.03	-0.31	0.26
Convicted and in treatment	4	0.14	-0.94	1.23
Probationed and in treatment	1	-0.44*	-0.78	-0.10
Remanded	1	-0.30*	-0.53	-0.06
Remanded and/or convicted	3	0.19	-0.22	0.59
Remanded and/or convicted and in treatment	2	0.90**	0.78	1.03
<i>Type of crime</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Gender violence	1	-1.54**	-1.76	-1.31
Homicidal sex offense	1	-0.20	-0.64	0.24
Mixed	6	0.43	0.00	0.87
Partner violence	2	-0.64	-1.33	0.05
Sex offenders	5	0.11	-0.26	0.47
<i>Source of reference scores</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
General norms	20	0.03	-0.22	0.28
Original article	1	0.85**	0.54	1.15

Note. CI = Confidence interval. UL = Upper limit. LL = Lower limit.

\* =  $p < .05$ . \*\* =  $p < .01$

## Impulsivity

The analyses for self-reported impulsivity found no significant overall differences between forensic populations and reference groups ( $g = -.16$ , 95% CI  $[-.71, .39]$ ,  $p = .56$ ), for the eight corresponding studies ( $n = 1,664$ ). Results showed between-studies heterogeneity ( $I^2 = 98.144$ ,  $Q = 484.83$ ,  $p < .001$ ). The eight studies used the Barratt Impulsiveness Scale (BIS) version 11, the BIS version 10, the Emotionality, Activity, Sociability, and Impulsivity (EASI) scale, the Eysenck Impulsiveness Scale (EIS), and the SNAP. Moderator analyses showed significant effects for the assessment instrument used ( $Q = 47.21$ ,  $df = 4$ ,  $p < .001$ ) and sample origin ( $Q = 23.95$ ,  $df = 3$ ,  $p < .001$ ). Table 5 shows effect sizes per

instrument used and for each country of origin, indicating significantly different levels of self-reported impulsivity on the EASI and in German and Spanish samples.

Inspection of frequencies revealed that all subscales were named 'impulsivity'. Due to this lack of variance, no moderator analysis was carried out on this possible moderator variable. A moderator analysis of judicial phase showed a small significant moderation effect ( $Q = 343.62$ ,  $df = 4$ ,  $p < .001$ ;  $g = -0.16$ , 95% CI [-0.25, -0.07]), indicating significantly different levels of self-reported impulsivity between forensic populations and reference groups. Only one out of the eight studies featured both forensic scores and reference scores of a non-incarcerated sample. Results of the moderator analysis showed a small significant moderation effect ( $Q = 8.02$ ,  $df = 1$ ,  $p < .001$ ;  $g = 0.38$ , 95% CI [0.03, 0.74]) with effect sizes indicating different levels of self-reported impulsivity in forensic populations when compared to simultaneously created reference scores of a non-incarcerated sample. Regarding gender, results showed no significant effect on self-reported levels of impulsivity ( $Q = 1.64$ ,  $df = 1$ ,  $p = .20$ ;  $g = 0.09$ , 95% CI [-0.28, 0.47]). A moderator analysis for type of crime revealed a moderately significant effect ( $Q = 18.22$ ,  $df = 3$ ,  $p < .001$ ;  $g = 0.51$ , 95% CI [0.38, 0.63]). It also showed that forensic populations who were incarcerated for capital crimes and a mix of various categories of felonies, misdemeanors, and violations display different levels of self-reported impulsivity, when compared to non-forensic populations. Type of crime was not specified in four the studies focusing on impulsivity. All effect sizes are shown in Table 5.

Meta-regression analysis on mean age of the forensic sample showed a small significant effect (Coeff. =  $-0.15$ ,  $p = .04$ ), indicating that levels of self-reported impulsivity decrease with age. Year of study showed no significant effect (Coeff. =  $-0.07$ ,  $p = .24$ ). Inspection of the funnel plot for publication bias indicated the unlikelihood of a meaningful publication bias.

### Quality of studies

Methodological quality of the included studies was assessed independently by two of the authors (MS and MLM). Agreement among assessments (performed in STATA version 13; StataCorp, 2013) proved to be good at total NOS quality score level (73% agreement, Cohen's  $K = 0.57$ ,  $SE = 0.10$ ,  $p < .001$ ) and at item level (91% agreement,  $K = 0.84$ ,  $SE = 0.04$ ,  $p < .001$ ). Methodological quality of the studies was not associated with outcome at the level of statistical significance (Pearson's  $r = .12$ ,  $p = .20$ , Spearman's  $\rho = .13$ ,  $p = .19$ ).

## Discussion

The current study is, to the authors' knowledge, the first comprehensive meta-analytic review of self-reported severe variants of common personality dimensions – anger, aggression, hostility, antisocial traits, psychopathy, and impulsivity – carried out in forensic populations and compared to non-forensic norm or reference groups. This study found no overall differences in self-reported levels of anger, aggression, hostility, or impulsivity between the general or healthy population and forensic samples. It did find varying levels – both low and high – of self-reported anger, aggression, and hostility scores depending on what instrument or subscale was used and significantly decreasing levels of self-reported impulsivity with age. Self-reported antisocial and psychopathic features were significantly and substantially higher in forensic samples than in reference groups. Simultaneously, levels of these personality traits also varied per instrument and subscale. Levels of all studied personality traits also often varied with country of origin, gender, and judicial phase.

**Table 5.** *Impulsivity: Results of moderator analyses (8 studies)*

Instrument name	No. of studies	Hedge's <i>g</i>	95% CI LL	95% CI UL
BIS-10	2	0.30	-0.71	1.31
BIS-11	3	-0.27	-1.33	0.79
EASI	1	-1.41**	-1.81	-1.01
EIS	1	0.19	-0.05	0.42
SNAP	1	-0.14	-0.32	0.03
<i>Sample origin</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
US	4	-0.18	0.34	1.18
Italy (30%)	2	-0.58	-1.34	0.19
Germany	1	0.76**	-0.98	-0.21
Spain	1	-0.60**	-0.78	0.43
<i>Gender</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Female	3	0.23	-0.20	0.66
Male	5	-0.34	-1.11	0.42
<i>Judicial phase</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Convicted	4	-0.49	-1.11	0.12
Convicted and in treatment	1	0.76**	0.34	1.18
Remanded and/or convicted	1	0.19	-0.05	0.42
Remanded and/or convicted and in treatment	1	0.56**	0.42	0.70
<i>Type of crime</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Capital crimes	1	0.76**	0.34	1.18
Mixed crimes (felonies, misdemeanors, violations)	1	0.56**	0.42	0.70
Partner violence	2	-0.19	-0.95	0.57
<i>Source of reference scores</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
General norms	7	-0.27	-0.84	0.31
Original article	1	0.79**	0.34	1.24

Note. CI = Confidence interval. LL = Lower limit. UL = Upper limit.

\* =  $p < .05$ . \*\* =  $p < .01$

Results of moderation analyses suggest that a number of factors may influence self-reported severe variants of common personality dimensions, including assessment instrument and subscale used, country of sample origin, gender, judicial phase, type of crime, and source of the control or reference scores for self-reported levels of antisocial and psychopathic traits. However, due to the small number of studies in some of the groups, interpreting these results must be met with caution.

Consistent with the present study, previous research shows that self-reports of impulsivity as well as sensation-seeking and risk-taking decrease with age (Arnett, 1994; Galvan, Hare, Voss, Glover, & Casey, 2007; Leshem & Glicksohn, 2007; Steinberg, Albert, Cauffman, Banich, Graham, & Woolard, 2008). Also, studies have shown gender differences (Costa, Terracciano, & McCrae, 2001; Del Giudice, Booth, Irwing, 2012; Weisberg, DeYoung, & Hirsh, 2011) as well as cross-cultural differences in personality traits, such as higher mean scores on the MMPI-2 in Israeli samples than in American samples (Almagor & Nevo, 1996).

In the current study, levels of self-reported antisocial and psychopathic features were higher in forensic samples than in reference groups. This effect appeared to be especially large in studies that used the MMPI (54%) compared to studies using other instruments. Further inspection of this self-report personality assessment instrument revealed that the Psychopathic deviate (Pd) scale of the MMPI does not directly assess psychopathic or antisocial features but is made up of five rather heterogeneous subscales. These not only measure the antisocial constructs of social imperturbability and authority conflict but also problems with interpersonal relationships and impulse control, as well as social and self-alienation (Butcher, Hass, Greene, & Nelson, 2015). High scores on the Pd scale in the current study could, in theory, be due to other issues pertaining to social maladjustment. Unfortunately, the majority of the articles used in the current study only mentioned scores on the Pd scale as a whole. Therefore, precise scores per Pd subscale could not be investigated further to further analyze which subscale(s) in particular contributed to the high Pd score.

When considering the high levels of self-reported antisocial and psychopathic features found in the current study, one might wonder which came first: high levels of these traits or a prison sentence? Perhaps these traits worsen the longer an individual is imprisoned? This was in fact discussed by Adams (1976) in one of the oldest articles used in the current meta-analysis, when considering recidivism and high scores on the MMPI's psychopathic deviate scale: does each imprisonment only serve to worsen antisocial attitudes and continue the recidivistic cycle? Findings by Osberg and Poland (2001) also indicate that specific characteristics of psychopathy are related to history of previous crime in a sample of inmates in a maximum-security prison: problems with authority, self-alienation, and familial discord. A range of research, however, strongly supports the association between the antisocial lifestyle that is characteristic of antisocial personality disorder and psychopathy and a high rate of imprisonment (including Black, Gunter, Loveless, Allen, & Sieleni, 2010; Fazel & Danesh, 2002; Hare, 2006; Neumann & Hare, 2008). It could be the case that higher levels of antisocial and psychopathic traits, and consequently a more profound antisocial lifestyle, lead to higher prison sentences for offenders.

Previous studies have found under(self-)reporting of aggression and hostility (Hornsveld, Muris, Kraaimaat, & Meesters, 2009) and overall personality pathology (Spaans et al., 2015) in forensic samples. The lack of significant differences in levels of anger, aggression, hostility, or impulsivity between forensic and non-forensic populations in the current study could be due to the use of self-report instruments in forensic settings. Several authors have voiced concerns about the use of self-report data within forensic samples (Edens, 2009; Milton et al., 2005; Spaans et al., 2009) as they may yield underestimations of the actual level of forensically relevant personality traits due to their sensitivity to bias from a social desirable response tendency in prison inmates. This tendency to give positive self-descriptions (Paulhus, 2002) includes intentional positive impression management and

faking good (presenting oneself in a positive light) and unintentional self-deception (Ray, Hall, Poythress, Rivera-Hudson, & Lilienfeld, 2013) on items that clearly describe a negative trait. Possible aims are to gain advantages such as a lower prison sentence or to avoid enforced treatment. The lower than expected levels of anger, aggression, hostility, and impulsivity in forensic samples could in theory be due to this tendency.

On the other hand, previous research comparing self-report assessment with interview or observer-rated methods to establish their convergence for the diagnosis of personality disorder has yielded conflicting results about antisocial and psychopathic personality traits. De Ruiter and Greeven (2000), for example, found that Cluster B personality disorder was underreported on a self-report instrument compared to interview methods in a forensic sample. Hilderbrand and De Ruiter (2004) concluded that this was most likely due to the lack of self-awareness of any symptoms and defensiveness that is inherent to Cluster B personality disorders. Zimmerman and Coryell (1990) and Blackburn, Donnelly, Logan, and Renwick (2004) also found underreporting of passive-aggressive and antisocial personality disorder, respectively, in self-report instruments in a forensic sample. These findings do not concur with the present finding of higher levels of antisocial and psychopathic features in forensic samples and regrettably bring us no closer to an explanation of the current findings. It could be the case that lack of insight into one's own symptoms and a social desirable response tendency diminish differences in levels of self-reported severe variants of common personality dimensions between forensic and non-forensic groups, and that only large differences 'survive' this overall diminishing effect. This would explain the significantly higher levels of self-reported antisocial and psychopathic features in forensic samples, despite the above-mentioned response tendencies.

The results of the current study add to existing doubts about the validity of the self-report method for assessing deviant or disruptive personality traits. The differences per instrument used to assess self-reported anger, aggression, and hostility indicate the importance of finding a universal instrument and terminology that is suitable for the forensic field. Divergent results in forensic samples concerning distorted response styles such as positive impression management, especially in the assessment of antisocial and psychopathic personality traits, indicate the need for further research to determine the levels of severe variants of common personality traits in forensic samples. It might be the case that different kinds of deviant response styles apply to different personality traits, whereas it has also been found to be related to possible (legal) consequences (Cima, Pantus, & Dams, 2007; Walters, 1988). Further research should be conducted by applying a range of self-report instruments, standardized clinical interviews, and observer-rated assessment methods to examine if the personality traits anger, aggression, hostility, antisocial traits, psychopathy, and impulsivity are all forensically relevant if self-report methods do not uncover all of them in forensic populations. It is important to determine the most appropriate method of reliably assessing levels of the various forensically relevant personality traits in forensic samples, while keeping in mind response styles, such as positive impression management, that may distort results.

### **Limitations**

A limitation of the current study is that only five out of 39 studies featured means and standard deviations of self-reported levels of severe variants of common personality traits belonging to non-imprisoned samples along with those of a forensic population. This led to the necessity of consulting general normative scores in order to compare forensic samples to reference samples. Even though year of study, gender, and country of origin were matched as much as possible, results may have been

influenced by the use of these reference scores instead of scores from the same sample. In some cases where the normative scores could not be accessed through several channels, other samples than normative samples were consulted.

The three overarching categories of personality traits, in which results from 32 different subscales or constructs were grouped, were closely based on research by Miller and Lynam (2001; 2003; 2015), Miller and colleagues (2001), and Widiger and Costa (2012), with one category representing general features of antisocial and psychopathic behavior and the others representing agreeableness and conscientiousness, respectively. However, according to Krueger, Derringer, Markon, Watson, and Skodol (2012), some personality traits clearly belong to only one domain while others, may share features of more than one domain. For example, in the current study, dominance was included in the Antisocial/Psychopathy category. High dominance was previously found to be a characteristic of aggressive and antisocial individuals (Dolan & Blackburn, 2006; Doyle & Dolan, 2006). Perhaps if dominance had been placed in the Anger/Aggression/ Hostility category, results may have been different. Edens (2009), in a study included in the current meta-analysis, reported higher levels of self-reported dominance in prison inmates than both community and clinical normative samples.

## Conclusion

The current study suggests that forensic professionals should be cautious of the use of self-report instruments to determine levels of severe variants of common personality traits, as results may be prone to bias due to intentional impression management or unintentional self-deception. They should rather base conclusions on a combination of different assessment methods.



Chapter Three

***The DAPP-SF  
as a screener for  
personality disorder  
in a forensic setting***

## Abstract

Studies on the Dimensional Assessment of Personality Pathology-Short Form (DAPP-SF) have shown its ability to identify treatment-seeking patients with personality disorders. The present study focuses on its screening potential for personality disorder in 89 criminal suspects (77 men, 12 women; mean age 37.0 years) undergoing residential pre-trial psychological assessments in a high-security setting. It was expected that Structured Interview for DSM-IV Personality (SIDP-IV) criteria met for personality disorder(s) would be associated with higher DAPP-SF scores. A floor effect was found in DAPP-SF scores: the forensic population reported less personality pathology than the general population. Only moderate associations between DAPP-SF and SIDP-IV outcome were found. ROC analysis showed that some DAPP-SF subscales did not exceed chance level in their ability to screen for personality disorders. It is concluded that the DAPP-SF has limited usefulness as a screener for personality disorders in a forensic pre-trial setting. Alternative forensic screening instruments are presented.<sup>2</sup>

## Introduction

For an efficient assessment of personality psychopathology, a two-step approach is recommended (Widiger & Samuel, 2005). This approach entails first administering a self-report questionnaire to screen for the potential presence of personality disorders, followed by a standardized (semi-) structured diagnostic interview to verify the presence of the disorder, such as the Structured Interview for DSM-IV Personality (SIDP-IV; Pfohl, Blum, & Zimmerman, 2006). If the screening questionnaire were quick, efficient, and accurate, the amount of time required to confirm the presence or absence of a diagnosis would be shorter.

The Dimensional Assessment of Personality Pathology – Short Form (DAPP-SF; Van Kampen et al., 2008) meets the need for a concise measure to screen for personality pathology, as it takes on average only 20 minutes to complete. Studies have shown its reliability and validity in the general population as well as in patients seeking treatment for personality disorders (Van Kampen et al., 2008) and mood, anxiety, and somatoform disorders (De Beurs et al., 2009). A further study by De Beurs and colleagues (2010) concluded that the DAPP-SF was able to distinguish patients with personality disorders (ascertained by the SIDP-IV) from the general population.

The present study followed the example of De Beurs and colleagues (2010), but this time in a forensic sample, and focused on how well the DAPP-SF can screen for personality disorder in criminal suspects undergoing pre-trial psychological assessments in a high-security observation clinic, and whether it could determine correctly which suspects should and should not undergo the SIDP-IV interview in the two-step process proposed above. The outcome of the SIDP-IV was used to determine formal presence or absence of personality disorder and was compared to scores on the subscales of the DAPP-SF. It was expected that when SIDP-IV criteria were met for one or more personality disorders, the same individual would present with higher scores on subscales of the DAPP-SF.

### Pre-trial assessment in the Netherlands

Previous studies examining the screening capacity of the DAPP-SF have used treatment-seeking patients. In contrast, the present study investigates the utility of the DAPP-SF in screening for personality disorders among individuals undergoing mandated forensic pre-trial evaluation. Article

.....  
<sup>2</sup> Spaans, M., De Beurs, E., Rinne, T., & Spinhoven, P. (2015). The DAPP-SF as a screener for personality disorder in a forensic setting. *Journal of Personality Assessment*, 97(2), 172-181.

39 of the Dutch Criminal Law Code states: “Not punishable is he who commits a crime, that he cannot be held responsible for due to mental retardation or pathological impairment of his mental abilities”. In Dutch practice, this mental retardation or pathological impairment includes personality pathology (Barendregt, Muller, et al., 2008) and the offender can be ordered to undergo forensic psychiatric treatment. A person who is only partially responsible is considered partially punishable. Courts can order suspects of crimes that carry a penalty of four or more years of incarceration according to the Dutch Criminal Law Code to undergo a residential pre-trial psychiatric assessment to determine the presence of such defects or impairments at the time of the crime.

Depending on the severity of the disorder and the causality between the personality disorder and the crime, the criminal responsibility can be categorized in one of five degrees: full responsibility, slightly diminished responsibility, diminished responsibility, strongly diminished responsibility, and total lack of criminal responsibility. Each degree is translated by the judge into a different form of punishment. Fully responsible and slightly diminished responsible offenders will receive a prison sentence only. Offenders that are found fully not responsible, on the basis of a major psychiatric disorder that caused the crime, are discharged from punishment and ordered to undergo treatment in a high security forensic psychiatric clinic. Diminished and severely diminished offenders with a high risk of recidivism are usually given a shortened prison sentence, followed by imposed forensic treatment in a high security forensic psychiatric clinic that aims to treat the criminogenic defects or impairments in order to reduce the risk of re-offending (Koenraadt et al., 2007). An offender is released only when the risk of re-offending is sufficiently diminished.

The presence of personality disorders has somewhat of a unique role in determining criminal responsibility in the Netherlands compared to most other jurisdictions, where the presence of a personality disorder is not considered sufficient grounds for diminished responsibility or criminal insanity. Earlier research on Dutch forensic pre-trial assessments has shown that 78% of individuals with diminished criminal responsibility had a personality disorder, which was in turn associated with advice for enforced forensic treatment (Spaans, Barendregt, Haan, Nijman, & De Beurs, 2011).

## Method

### **Assessment site**

All participants were admitted to the Pieter Baan Center (PBC) in Utrecht (the Netherlands) for residential pre-trial criminal responsibility assessment and recidivism risk analysis. The PBC is the official forensic psychiatric observation clinic of the Dutch Ministry of Justice’s Netherlands Institute of Forensic Psychiatry and Psychology (NIFP), and has the legal status of a house of detention. The PBC administers close to 90% of all inpatient forensic assessments in the Netherlands, around 215 per year. These assessments cover 5% of all forensic pre-trial evaluations on adults carried out by the NIFP; the remaining 95% of forensic evaluations (roughly 4,200 per year) take place in a non-specialized forensic setting (usually in a regular house of detention). Possible reasons for the court to order a thorough inpatient assessment of defendants include the severity of the crime, the severity of the assumed psychopathology, the maximum-security level within the PBC, and potential societal disturbance or media attention associated with the defendant’s case. As a result, the population of the PBC covers the more severe criminological and psychiatric cases and cannot be seen as representative for the entire forensic population whose mental status is assessed.

All defendants are evaluated during a seven-week period by a multidisciplinary team consisting of a psychiatrist, a psychologist, two social workers, and a lawyer who supervises the assessment process along with a second psychiatrist. One of the social workers investigates the life history and social background of the defendant through interviews with informants such as family members; the other is a supervisor on the defendant's ward whose task is to observe and describe the activities and behavior of defendant during his or her stay in the institution. The psychologist and psychiatrist carry the final responsibility in the team for consensus in the PBC's conclusion in the final report to the judge concerning DSM-IV diagnoses, if any, and criminal responsibility (based on clinical judgment underpinned by structured instruments). The latter two experts also advise the court whether forensic treatment of the defendants is indicated if convicted of the charge.

## Participants

The data collected in this study stemmed from a subset of all 839 criminal suspects who were admitted to the PBC between October 2007 and October 2011. Of these 839 suspects, 266 (31.7%) cooperated with the voluntary psychological testing procedure, which includes the administration of a variety of instruments. The study sample consisted of all 143 criminal suspects who had completed a DAPP-SF self-report questionnaire between October 2007 and October 2011. Of this total sample, 89 suspects had also completed the SIDP-IV interview during that time, and 54 suspects completed only the DAPP-SF. The data for this study was collected by test assistants and clinical psychologists who had completed a training session in administration and scoring of both instruments prior to data collection. Training for the SIDP-IV was provided by an academic and clinical expert, in the form of an in-company session lasting three hours. Reaching consensus on symptom criteria and weighing and interpreting results were trained and discussed. Inter-rater data on SIDP-IV administrations in the PBC is not available, as raters were trained to conduct independent and individual interviews. Any questions test assistants had in scoring criteria in individual interviews were discussed with the corresponding multidisciplinary team. The sample of 89 suspects who completed both the DAPP-SF and SIDP-IV interview consisted of 77 men (86.5%) and 12 (13.5%) women. The mean age at the time of the forensic assessment was 37.0 years ( $SD = 11.6$  years; range 18.2-66.9 years). Intelligence of 77 participants (86.5%) was assessed with the Dutch version of the Wechsler Adult Intelligence Scale-III (WAIS-III; Uterwijk, 2000). Ten participants (11.2%) were assessed using the Dutch version of the Kaufman Adult Intelligence Scale (KAIT; Mulder, Dekker, & Dekker, 2004) because a WAIS-III had been administered elsewhere recently. One participant was assessed using the Multicultural Capacity Test (MCT; Bleichrodt & Van en Berg, 1999) because of insufficient knowledge of the Dutch language. Mean IQ total score was 95.0 ( $SD = 14.2$ ). Of the 89 participants, six (6.72%) had an above average intelligence (a score between 110-120) and five (5.6%) had a high intelligence (a score higher than of 120). Average intelligence (a score between 90-110) was recorded for 44 (49.4%) participants, and 22 (24.7%) had a below average intelligence (a score between 80-90). Eleven participants (12.4%) were assessed at the level of borderline intellectual functioning (a score lower than 80) but still considered capable of undergoing psychological testing with the DAPP-SF and/or SIDP-IV by the forensic psychologist. For one participant (1.1%) the intelligence level was unknown. There were no significant differences in IQ scores between men and women.

Of the 89 suspects, 32 had been accused of (attempted) murder, followed by (attempted) sex offenses ( $n = 20$ ) and (attempted) manslaughter ( $n = 16$ ). Seven suspects had been accused of (attempted) armed robbery, four of arson, and three of extortion. The remaining ten suspects had been accused of (attempted) grievous bodily harm, bodily harm, kidnapping, threatening violence, and neonaticide.

## Measures

### Dimensional Assessment of Personality Pathology-Short Form

The DAPP-SF is a self-report questionnaire that assesses the presence and severity of personality pathology. It has 126 items measuring personal preferences and behavior and is the shortened version of the Dimensional Assessment of Personality Pathology – Basic Questionnaire (DAPP-BQ; Livesley & Jackson, 2006) which has 290 items. The DAPP-SF is comprised of 18 personality dimensions (Submissiveness, Cognitive Distortion, Identity Problems, Affective Lability, Stimulus Seeking, Compulsivity, Restricted Expression, Callousness, Oppositionality, Intimacy Problems, Rejection, Anxiousness Conduct Problems, Suspiciousness, Social Avoidance, Narcissism, Insecure Attachment, and Self-harm) and four second-order factors (Emotional Dysregulation, Dissocial Behavior, Inhibition, and Compulsivity). Each item is rated on a five-point Likert scale.

Psychometric analysis of the DAPP-SF has revealed sufficient reliability with alpha coefficients ranging from 0.78 to 0.89 with a mean value of 0.84, as well as construct validity and congruent factor structure (Tucker's congruence coefficients ranging from 0.89 to 1.00) in the general population and in patients seeking treatment for personality disorders (Van Kampen et al., 2008). The same was also found in patients seeking treatment for mood, anxiety, and somatoform disorders (De Beurs et al., 2009).

Van Kampen and colleagues (2008) compared DAPP-SF scores of three groups of respondents (patients seeking treatment for personality disorders, patients seeking treatment for depressive, anxiety, and somatoform disorders, and a population-based sample from general practitioners' patients) to scores on the SIDP-IV and the Five-Dimensional Personality Test (5DPT; Coolidge, Segal, Cahill, & Archuleta, 2008). Positive correlations were obtained between the DAPP-SF second-order factors Emotional Dysregulation, Dissocial Behavior, and Compulsivity and the 5DPT dimensions Neuroticism, Insensitivity and Orderliness. A negative correlation was obtained between Inhibition and Extraversion. Van Kampen and colleagues (2008) concluded that all indices of convergent and divergent validity were satisfactory. A high score on a personality dimension or second-order factor indicates a high probability of a personality disorder. For example, the DAPP-SF manual (Van Kampen & De Beurs, 2009) gives cut-off scores for the dimension Identity Problems, Anxiousness, and Social Avoidance that provide optimal specificity to distinguish personality-disordered respondents. Respondents who score above these cut-off values most likely meet the criteria for a personality disorder and, therefore, qualify for further clinical assessment using a standardized diagnostic interview. As mentioned earlier, De Beurs and colleagues (2010) also found support for the screening potential of the DAPP-SF in patients with personality disorders from the general population. For instance, a score of 180 on the Emotional Dysregulation scale had an AUC of 0.87 and resulted in a good sensitivity of 0.90 with a still acceptable specificity of 0.76.

### Structured Interview for DSM-IV Personality

The SIDP-IV comprises ten sections, covering different life areas such as activities and interests, work, relationships, emotions, self-perception, stress, and anger. The interview follows a natural course on these and other topics so that questions about DSM-IV criteria are not as predictable as in other instruments. For each diagnosis of personality disorder (Paranoid, Schizoid, Schizotypal, Antisocial, Borderline, Histrionic, Narcissistic, Avoidant, Dependent, Obsessive-Compulsive, Self-Defeating, Depressive, and Negativistic) a dichotomous score indicates the diagnostic status and a dimensional score represents the severity of the disorder. Research by Damen, De Jong and Van der Kroft (2004) has shown that the Dutch version of the SIDP-IV has good inter-rater reliability.

## Data analysis

Cronbach's alpha coefficients were calculated for the DAPP-SF dimension and second-order factor scores to determine the internal consistency for the study sample. Spearman's rank-order correlations were calculated between DAPP-SF dimension and second-order factor scores and the number of PD diagnoses according to the SIDP-IV. Pearson's correlations were calculated between DAPP-SF dimension and second-order factor scores and the number of criteria met on each SIDP-IV PD. Independent samples *t*-tests were conducted to compare mean DAPP-SF dimension and second-order factor scores in the study sample with each other, as well as with those of the general population and the presence or absence of a PD diagnosis according to the SIDP-IV. A Receiver Operating Characteristic (ROC) analysis was carried out for DAPP-SF dimension and second-order factor scores and the presence or absence of a PD diagnosis according to the SIDP-IV.

## Results

### DAPP-SF descriptives

Table 1 shows descriptive statistics for personality dimension scores and second-order factor scores on the DAPP-SF for the sample of participants who underwent both the DAPP-SF and the SIDP-IV interview ( $n = 89$ ). Also shown are mean dimension and second-order factor scores and standard deviations from the Dutch general population studied by De Beurs and colleagues (2010) and effect sizes for the comparison of the two populations. Cronbach's alpha coefficients ranged from 0.75 for Conduct Problems to 0.91 for Social Avoidance and Self-harm. Cronbach's alpha coefficients for second-order factor scores ranged from 0.73 for Inhibition to 0.98 for Emotional Dysregulation. The internal consistency found in this forensic sample does not differ greatly from that in the general population. Comparison of the mean scores of the study population and the general population, using a *t*-test for independent samples, found that scores are significantly lower for the forensic participants compared to the general population on all dimensions except Identity Problems, Intimacy Problems, Conduct Problems, Insecure Attachment, and Self-Harm and all second-order factors except Inhibition ( $d = 0.27 - 0.72$ ). Effect sizes varied from small for Insecure Attachment and Self-Harm to large for Intimacy Problems and Inhibition. Overall, the study population reported less personality pathology than the general population.

### Comparing samples with and without a SIDP-IV

To investigate potential selection bias in the study population, DAPP-SF scores of the 54 participants who did not undergo a SIDP-IV interview were compared to the scores of the 89 participants who underwent both the DAPP-SF and the SIDP-IV interview. Means of all 22 personality dimension scores and second-order factor scores in both samples were compared using a *t*-test for independent samples, which yielded no significant differences between the DAPP-SF scores of participants with or without a SIDP-IV according to adjusted alpha levels of 0.005 (see Table 2).

**Table 1.** Descriptives of DAPP-SF personality dimension and second-order factor scores

	Study population (n = 89)			General population (n = 461)		t	Cohen's d
	$\alpha$	M	SD	M	SD		
<i>Personality dimensions</i>							
Submissiveness	.88	16.14	6.75	19.68	6.32	4.78	-.54
Cognitive distortion	.81	8.89	3.96	12.12	5.40	5.37	-.68
Identity problems	.86	11.16	5.29	12.18	5.64	1.58	-.19
Affective lability	.86	17.21	6.58	21.04	7.28	4.61	-.55
Stimulus seeking	.87	16.29	6.81	18.00	5.76	2.49	-.27
Compulsivity	.86	19.97	7.04	24.24	6.48	5.61	-.63
Callousness	.77	16.92	5.29	18.80	5.40	3.02	-.35
Restricted expression	.85	19.27	6.69	21.28	6.48	2.67	-.31
Oppositionality	.86	19.01	7.15	23.10	7.20	4.91	-.57
Intimacy problems	.78	27.04	6.47	16.96	5.68	15.0	1.66
Rejection	.78	17.28	5.32	20.08	5.68	4.30	-.51
Anxiousness	.87	13.29	5.73	14.82	5.70	2.32	-.27
Conduct problems	.75	13.75	5.37	11.52	4.40	4.21	.45
Suspiciousness	.87	13.18	5.68	14.96	5.92	2.61	-.31
Social avoidance	.91	10.86	5.43	13.80	5.52	4.61	-.54
Narcissism	.82	14.99	5.65	18.72	6.16	5.30	-.63
Insecure attachment	.89	13.21	6.45	13.74	5.64	.79	-.09
Self-harm	.91	8.10	4.30	7.98	4.20	.25	.03
<i>Second-order factors</i>							
Emotional Dysregulation	.98	146.03	50.36	156.96	46.08	2.02	-.23
Inhibition	.73	43.96	8.18	32.32	7.52	13.18	1.48
Dissocial Behavior	.92	66.59	19.00	80.92	20.74	6.05	-.72
Compulsivity	.86	19.97	7.04	24.24	6.48	5.61	-.63

**Table 2.** Comparison of DAPP-SF personality dimension and second-order factor mean scores for participants with and without a SIDP-IV interview (n = 143)\*

	DAPP-SF and SIDP-IV (n = 89)		DAPP-SF only (n = 54)		t	Cohen's d
	M	SD	M	SD		
<i>Personality dimensions</i>						
Submissiveness	16.14	6.75	15.35	6.00	-.71	-.12
Cognitive distortion	8.89	3.96	8.53	3.50	-.56	-.09
Identity problems	11.16	5.29	11.20	5.41	.05	.01
Affective lability	17.21	6.58	17.05	6.74	-.14	-.02
Stimulus seeking	16.29	6.81	15.24	5.59	-.95	-.16
Compulsivity	19.97	7.04	22.17	6.87	1.82	.31
Callousness	16.92	5.29	17.58	5.81	.71	.12
Restricted expression	19.27	6.69	20.22	7.06	.81	.14
Oppositionality	19.01	7.15	18.00	6.89	-.91	-.15
Intimacy problems	27.04	6.47	26.71	7.24	-.28	-.05
Rejection	17.28	5.32	18.47	5.81	1.50	.25
Anxiousness	13.29	5.73	13.35	6.27	.07	.01
Conduct problems	13.75	5.37	13.37	5.31	-.42	-.07
Suspiciousness	13.18	5.68	14.90	7.02	1.61	.27
Social avoidance	10.86	5.43	10.66	5.76	-.20	-.03
Narcissism	14.99	5.65	15.46	5.48	.49	.08
Insecure attachment	13.21	6.45	11.83	5.43	-1.31	-.22
Self-harm	8.10	4.30	8.09	4.18	-.01	.00
<i>Second-order factors</i>						
Emotional Dysregulation	146.03	50.36	144.34	49.92	-.20	-.03
Inhibition	43.96	8.18	44.30	9.38	.23	.04
Dissocial Behavior	66.59	19.00	67.30	18.07	.22	.04
Compulsivity	19.97	7.04	22.17	6.87	1.82	.31

### **SIDP-IV descriptives**

In the sample of 89 participants who underwent both the DAPP-SF and the SIDP-IV interview, a total of 45 personality disorders were diagnosed in 32 individuals. In 22 cases (24.7%) one personality disorder was diagnosed. In seven cases (7.9%) two separate personality disorders were diagnosed and in three cases (3.4%) three personality disorders were diagnosed in the same person according to the SIDP-IV. Cluster B personality disorders were the most prevalent with a total of 30 diagnoses: 16 for Antisocial Personality Disorder, nine for Narcissistic PD, five for Borderline PD, and one for Histrionic PD. There were four diagnoses for Negativistic PD, four for Paranoid PD, one for Schizoid PD and one for Depressive PD. Cluster C was the least prevalent with only four diagnoses: two for Avoidant PD, one for Dependent PD, and one for Obsessive-Compulsive PD. There were no diagnoses for Schizotypal or Self-Defeating PD.

### **Comparison of SIDP-IV diagnoses and expert opinions**

To assess the legitimacy of using the SIDP-IV classification as a criterion for personality disorder (PD) diagnosis in the present study, the SIDP-IV outcome was compared to the clinical opinions of the forensic expert teams in the PBC reports for all 89 participants who had completed both a DAPP-SF and a SIDP-IV interview. When reaching their final diagnosis, all expert teams integrated the SIDP-IV interview outcome with additional observation information from the suspect's seven-week admission and incorporate relevant information from the suspect's life history.

In 20 cases (22.5%) the experts' conclusion cited a PD Not Otherwise Specified, a classification that is officially not a possibility on the SIDP-IV and, therefore, not entirely comparable. Relevant traits of the PD indicated by the experts, however, could be found upon inspection on SIDP-IV item level for these 20 cases. This indicates that the experts' conclusions and the SIDP-IV outcome were to a large extent in accordance with each other.

Of the remaining 69 cases for which an exact comparison was possible, complete agreement was found for 50 participants (72.5%). In 13 of the remaining 19 cases the SIDP-IV found no PD while the forensic experts did (having incorporated information from the suspect's seven-week stay on the ward and important life history information to diagnose a PD). In six cases the SIDP-IV found a PD while the forensic expert team did not (acknowledging the many traits that were present of the PD, but finding no impairments in daily functioning to warrant a PD diagnosis or diagnosing a different disorder such as Autism Spectrum Disorder and Delusional Disorder). These cases show that when coming to a PD diagnosis, experts incorporate traits found on the SIDP-IV with their own observations, information from the PBC ward, and important life history information on impairments in daily functioning in the present and the past. Finding only 19 out of the comparable 69 six cases (27.5%) in which the SIDP-IV outcome did not converge with the forensic expert team's clinical findings gives an extra boost of confidence in the validity of the SIDP-IV classification and warranted its use as a criterion for PD diagnosis (as opposed to the teams' opinions) in this study.

### **The DAPP-SF as a screener for PD**

To investigate the ability of the DAPP-SF to discriminate between participants with and without one or more personality disorders, the DAPP-SF scores were compared for individuals with and without a PD according to the SIDP-IV ( $n = 89$ ). Independent samples  $t$ -tests (see Table 3) found no differences in personality dimension scores and second-order factor scores on the DAPP-SF for participants with and without a PD according to the SIDP-IV for adjusted alpha levels of .005. Most effect sizes were negligible or small, except for Stimulus Seeking, Restricted Expression, and Dissocial Behavior with

sizes between .47 and .49. Table 4 shows Spearman's rank-order correlations between the DAPP-SF personality dimension scores and second-order factor scores and the number of PD diagnoses (ranging from 0 to 3) for the entire study population. No significant associations were found according to the adjusted alpha levels of .005 and all associations were negligible as none were higher than .25. Correlational analyses for the three separate clusters of personality disorder were not possible due to the small number of participants in each cluster.

**Table 3.** Comparison of DAPP-SF personality dimension and second-order factor mean scores for participants with and without a PD according to the SIDP-IV interview (n = 89)\*

	PD present (n = 32)		PD absent (n = 57)		t	Cohen's d
	M	SD	M	SD		
<i>Personality dimensions</i>						
Submissiveness	15.46	6.46	16.53	6.93	.72	.15
Cognitive distortion	9.10	3.55	8.78	4.20	-.37	-.08
Identity problems	11.91	5.15	10.74	5.36	-1.00	-.21
Affective lability	17.80	6.68	16.88	6.56	-.63	-.14
Stimulus seeking	18.38	7.28	15.12	6.30	-2.21	-.47
Compulsivity	20.60	7.22	19.62	6.98	-.62	-.13
Callousness	17.09	5.64	16.81	5.13	-.24	-.05
Restricted expression	21.31	6.54	18.12	6.56	-2.20	-.47
Oppositionality	20.22	7.56	18.33	6.89	-1.20	-.26
Intimacy problems	27.58	5.53	26.74	6.97	-.59	-.13
Rejection	17.74	5.00	17.03	5.51	-.60	-.13
Anxiousness	14.73	6.12	12.47	5.39	-1.81	-.39
Conduct problems	15.17	5.73	12.95	5.03	-1.91	-.41
Suspiciousness	13.13	5.67	13.21	5.74	.07	.02
Social avoidance	10.75	4.37	10.91	5.98	.14	.03
Narcissism	15.82	7.02	14.53	4.73	-1.03	-.22
Insecure attachment	13.81	6.37	12.87	6.52	-.66	-.14
Self-harm	8.50	4.87	7.88	3.97	-.65	-.14
<i>Second-order factors</i>						
Emotional Dysregulation	151.22	48.17	143.12	51.74	-.73	-.16
Inhibition	44.68	8.01	43.55	8.31	-.62	-.13
Dissocial Behavior	72.60	18.33	63.21	18.69	-2.29	-.49
Compulsivity	20.59	7.21	19.62	6.98	-.62	-.13

\* No significant values were found according to adjusted alpha levels of 0.005

**Table 4.** Spearman's rank-order correlations between DAPP-SF scores and number of PD diagnoses according to the SIDP-IV (n = 89)\*

Number of SIDP-IV PD diagnoses	
<i>Personality dimensions</i>	
Submissiveness	-.03
Cognitive distortion	.13
Identity problems	.19
Affective lability	.13
Stimulus seeking	.25
Compulsivity	.06
Callousness	.02
Restricted expression	.25
Oppositionality	.16
Intimacy problems	.04
Rejection	.09
Anxiousness	.22
Conduct problems	.17
Suspiciousness	.01
Social avoidance	.13
Narcissism	.02
Insecure attachment	.13
Self-harm	.11
<i>Second-order factors</i>	
Emotional Dysregulation	.15
Inhibition	.05
Dissocial Behavior	.27
Compulsivity	.06

\* No significant values were found according to adjusted alpha levels of 0.005

### SIDP-IV symptom profiles

The ability of the DAPP-SF to discriminate between participants with and without one or more personality disorders according to the SIDP-IV was investigated once more with Pearson's correlations for DAPP-SF subscale and second-order factor scores and the *number* of criteria met on each SIDP-IV PD. The latter can function as a profile of PD symptoms, which better suits the dimensional character of the DAPP-SF. Results are shown in Table 5. A number of significant correlations were found according to adjusted alpha levels of 0.005. Although the overall pattern of significant correlations showed some concurrence between DAPP-SF dimensions and the number of criteria met on each personality disorder according to the SIDP-IV – such as for Borderline, Histrionic, and Depressive Personality Disorder – and suggests that the dimensional approach to the SIDP-IV outcomes are in line with expectations, the correlation coefficients revealed primarily moderate associations (ranging from  $r = -.22$  to  $r = .48$ ).

**Table 5.** Pearson's correlations between DAPP-SF scores and number of criteria met per SIDP-IV PD (n = 89)

SIDP-IV personality disorders						
DAPP-SF	Schizoid	Borderline	Avoidant	Dependent	Obsessive-compulsive	Negativistic
<i>Personality Dimensions</i>						
Submissiveness	.29	.30*	.36*	.18	.08	-.07
Cognitive distortion	.05	.39*	.06	.09	-.05	-.11
Identity problems	.21	.48*	.32*	.15	.09	.01
Affective lability	.15	.45*	.24	.16	.11	.04
Stimulus seeking	.21	.44*	.11	-.06	-.01	.01
Compulsivity	-.12	-.14	.05	.04	.29	-.03
Callousness	.10	.09	-.02	-.02	.04	-.15
Restricted expression	.30*	.33*	.28	.07	.14	.02
Oppositionality	.20	.47*	.18	-.04	-.04	-.06
Intimacy problems	-.03	.05	-.01	-.05	-.01	-.12
Rejection	-.02	.09	-.13	.01	.10	-.10
Anxiousness	.17	.45*	.23	.16	.12	.09
Conduct problems	.08	.28	-.08	-.22	-.11	.11
Suspiciousness	.20	.11	.07	.15	.10	-.01
Social avoidance	.38*	.37*	.29	.16	.07	.04
Narcissism	-.07	.21	-.05	-.07	-.03	-.13
Insecure attachment	.06	.31*	.16	.20	.24	.08
Self-harm	.16	.38*	.36*	.13	-.02	.10
<i>Second-order factors</i>						
Emotional Dysregulation	.21	.45*	.26	.14	.08	-.00
Inhibition	.04	.10	-.02	-.05	.02	-.19
Dissocial Behavior	.20	.38*	.08	-.06	.04	.01
Compulsivity	-.12	-.14	.05	.04	.29	-.03

\* Correlation is significant at the 0.005 level (2-tailed)

SIDP-IV personality disorders							
DAPP-SF	Self-defeating	Paranoid	Schizotypal	Histrionic	Narcissistic	Depressive	Antisocial
<i>Personality Dimensions</i>							
Submissiveness	.09	.26	.26	.01	-.09	.36*	-.01
Cognitive distortion	.07	.17	-.01	.01	-.16	.12	.01
Identity problems	.13	.26	.32*	.13	-.06	.35*	.02
Affective lability	.19	.28	.22	.11	-.04	.32*	-.08
Stimulus seeking	-.04	.31*	.21	.11	-.01	.12	.34*
Compulsivity	-.06	.03	-.11	-.05	.09	.05	-.00
Callousness	-.07	.24	.03	-.01	.15	-.01	.14
Restricted expression	.06	.27	.26	.05	-.03	.28	.18
Oppositionality	.03	.23	.24	.12	-.02	.18	.20
Intimacy problems	.03	-.12	-.11	.13	.08	.00	.05
Rejection	-.10	.34*	-.10	.00	.24	-.10	.02
Anxiousness	.23	.23	.20	.05	-.05	.38*	.09
Conduct problems	-.14	.14	-.03	-.09	.14	-.03	.41*
Suspiciousness	.01	.40*	.14	-.09	.01	.21	-.01
Social avoidance	.13	.21	.30*	-.01	-.13	.37*	.05
Narcissism	-.16	.10	-.06	.05	-.05	-.05	.08
Insecure attachment	.19	.24	.03	-.02	-.09	.32*	.02
Self-harm	.14	.08	.32*	-.05	-.11	.30*	-.08
<i>Second-order factors</i>							
Emotional Dysregulation	.12	.29	.23	.04	-.08	.33*	.04
Inhibition	-.02	.06	-.07	.10	.16	-.01	.14
Dissocial Behavior	-.06	.34*	.13	.03	.09	.11	.31*
Compulsivity	-.06	.03	-.11	-.05	.09	.05	-.00

\* Correlation is significant at the 0.005 level (2-tailed)

**Table 6.** Cut-off scores for DAPP-SF dimensions and second-order factors with optimum sensitivity (n = 89)

	Cut-off score	Sensitivity	Specificity
<i>Personality dimensions</i>			
Submissiveness	11	.72	.28
Cognitive distortion	6	.69	.44
Identity problems	8	.72	.42
Affective lability	13	.75	.39
Stimulus seeking	12	.72	.38
Compulsivity	15	.72	.32
Callousness	13	.69	.30
Restricted expression	16	.75	.49
Oppositionality	15	.72	.42
Intimacy problems	23	.75	.25
Rejection	15	.63	.39
Anxiousness	10	.72	.46
Conduct problems	11	.69	.47
Suspiciousness	8	.81	.23
Social avoidance	7	.78	.44
Narcissism	12	.66	.46
Insecure attachment	9	.75	.35
Self-harm	6	.47	.67
<i>Second-order factors</i>			
Emotional Dysregulation	120	.72	.42
Inhibition	42	.59	.39
Dissocial Behavior	61	.72	.49
Compulsivity	18	.63	.54

### ROC analysis

A Receiver Operating Characteristics (ROC) analysis was carried out for all DAPP-SF dimensions and second-order factors and the presence/absence of SIDP-IV personality disorders to determine the ability of the DAPP-SF to discriminate between the 32 individuals with and 57 without a personality disorder. The AUC for the 22 scores ranged from .45 (Submissiveness) to .66 (Dissocial behavior). Of the personality dimensions, Restricted Expression had the highest AUC (.65). No asymptotic significance was less than  $p = .05$ . Table 6 shows the cut-off scores per dimension and per second-order factor for the study population that correspond with optimum sensitivity and acceptable specificity. As can be seen, the proposed cut-off of 61 for Dissocial Behavior, for example, has a sensitivity of .72 (28% of the true cases are missed) and a specificity of .49 (51% of the participants without a PD are false positives on the screening instrument).

## Discussion

Based on the results presented above it can be concluded that, for a two-staged assessment process in a forensic context, the DAPP-SF has limited usefulness as a screener for personality disorders in a sample of criminal suspects undergoing forensic pre-trial examination in a high-security observation clinic. Even though a few small to medium effect sizes were obtained that give some support for the use of the DAPP-SF as a screening instrument for personality disorders, inspection of the mean scores showed that the forensic population reported less personality pathology than the general population. This discourages the use of the DAPP-SF in forensic populations as the question arises whether the forensic population really does experience less personality pathology or whether the possibility that the participants dissimulate or display positive impression management when they fill in the self-report questionnaire in order to decrease their chances of enforced forensic treatment should be considered. A floor effect appears to have led to the inability of the DAPP-SF scores to discriminate the absence or presence of personality disorders according to the SIDP-IV within the study population, because scores of both groups were equally low. Spearman's rank-order and Pearson's correlations and independent samples *t*-tests between DAPP-SF scores and SIDP-IV PD diagnoses yielded only three moderate associations, insufficient to distinguish individuals with personality disorders. AUC results of the ROC analysis indicated that, in line with the other results, some of the DAPP-SF dimensions and second-order factors did not exceed chance level in their ability to screen for personality disorders in the study population.

Although convergence between self-report instruments and interview methods for establishing personality disorder was found to vary per disorder (Blackburn et al., 2004), Guy, Poythress, Douglas, Skeem and Edens (2008) found that self-report and interview measures of personality disorder were related most strongly at a dimensional level (i.e., the symptom count). This was also expected in the present study when the SIDP-IV results were approached in a dimensional way to match the dimensional nature of the DAPP-SF. However, the Spearman's rank-order and Pearson's correlation coefficients revealed, at best, moderate associations.

A short description of the Dutch forensic context may explain the possibility of dissimulation or positive impression management by the suspects. As mentioned earlier, Dutch courts can order those found guilty of crimes that carry a penalty of four or more years of incarceration to receive enforced forensic treatment if there is reduced criminal responsibility. Suspects undergoing pre-trial forensic assessment in the Netherlands have a keen interest in *avoiding* enforced forensic treatment. Courts will not end the enforced forensic treatment until the criminogenic aspects of the disorder have been treated and the risk of re-offending has been sufficiently reduced. Consequently, there is no way of predicting how long the individual will spend in the forensic clinic. A current trend is one of caution in ending treatments and releasing forensic patients, after incidents with former patients led to public upheaval. Consequently, the average length of stay in forensic psychiatric treatment has almost doubled over the past years, from 4.2 years in 1990 to 7.7 years in 2007 (Brand & Van Gemmert, 2009). This has caused more and more suspects to refuse to cooperate with pre-trial assessments to avoid enforced forensic treatment at all costs. Understandably, they prefer the chance of a longer prison sentence over additionally imposed mandatory treatment for an indefinite period of time. Many defendants refuse their cooperation during the pre-trial assessment and do not talk to the forensic experts or participate in any psychological interviews or questionnaires. Those who do cooperate are motivated to present themselves as psychologically healthy individuals.

Many suspects that do decide to undergo any psychological testing are thus likely inclined to dissimulate, to showing themselves in a better light to try and dispel forensic experts' suspicions of any kind of pathology. As the DAPP-SF is a self-report questionnaire, it leaves ample room for biased results due to the tendency to dissimulate described above.

Proof of dissimulation by positive impression management in a forensic context has been found by several authors (Ahlmeyer, Heil, McKee, & English, 2000; Gutheil 2003; McEwan, Davis, MacKenzie, & Mullen, 2009; Mills, Loza & Kroner, 2003). Gutheil (2003) asserts that inmates may be encouraged by their attorneys to present their symptoms in a certain, tactical way. Caruso, Benedek, Auble, and Bernet (2003) found proof of two types of dissimulators: intentional and unintentional. Intentional dissimulators were motivated by the preference of a defined prison term over an undefined term of hospitalization and the wish to avoid stigmatization, for example. Unintentional dissimulation was not a rational choice, but a genuine lack of knowledge or awareness of one's psychiatric disorder or symptoms. This would also lead to low self-reported personality pathology.

It is likely that the self-report nature of the DAPP-SF makes it unsuitable as a screener in a forensic pre-trial examination context. The present study and research by Spaans and colleagues (2009) suggest that although self-report instruments are of great value in individual use and case-finding of possible personality pathology, they may be less suited for the first step in the two-step approach in forensic populations. A number of authors advise against self-report instruments in forensic populations, unless they contain a measure for positive impression management or dissimulation (De Beurs & Barendregt, 2008; De Ruiter & Greeven, 2000; Edens & Ruiz, 2006; Mills et al., 2003). Even though research by McGrath, Mitchell, Kim and Hough (2010) has cast doubt upon the justification of the use of such response bias indicators, Eden and Ruiz (2006) found direct support for the validity scales of the Personality Assessment Inventory (PAI; Morey, 1991) in a correctional setting. Anderson, Sellbom, Wygant and Edens (2013) found support for validity scales, including one to identify dissimulation or positive impression management, of the Psychopathic Personality Inventory-Revised (PPI-R; Lilienfeld & Widows, 2005), a self-report instrument for personality traits associated with psychopathy that contains items intended not to evoke socially desirable responses.

Using self-report instruments that include response bias indicators such as the PAI and the PPI-R, or clinician-administered (semi-)structured interviews, or having close relatives or partners complete a measurement instrument as a proxy to the individual likely to fake good are alternatives to the DAPP-SF for diagnosing personality disorders in a forensic context.

Although positive impression management by the forensic population in the current study is considered the most likely explanation for the findings, it could not be formally assessed in the current study. An alternative explanation could be lack of statistical power given the small sample size of 89 participants with both DAPP-SF and SIDP-IV data. Another study limitation is related to the possible selection bias of including only cooperating, sufficiently Dutch-speaking, non-psychotic participants. A further limitation was the lack of data on inter-rater reliability for the PBC's forensic experts on the SIDP-IV, particularly in light of evidence that the inter-rater reliability of other well-known clinical instruments was lower in the (forensic) field than stated in the test manuals, such as the Psychopathy Checklist-Revised (PCL-R; Hare, 2003; Miller, Kimonis, Otto, Kline, & Wasserman, 2012; Murrie, Boccaccini, Johnson & Janke, 2008) - even when conducted by legally independent forensic experts (Sturup et al., 2013) - and STATIC-99 (Hanson & Thornton, 2000; Miller et al., 2012). The results suggesting the DAPP-SF's unsuitability as a screener might be influenced by the choice of outcome criterion. Nonetheless, while the incorporation of the SIDP-IV interview outcome in the clinician's

final diagnosis – and the clinician therefore not being blind to the SIDP-IV outcome – is a standard practice in the PBC that could not be altered for the present study, the 72.5% convergence between the SIDP-IV classification and the clinical diagnosis and the extensive way in which most dissimilarities between the two could be clarified by additional expert information strengthened our decision to use the SIDP-IV as the criterion for the presence of PD.

It is important to mention that even suspects who cooperate in their pre-trial assessment fear enforced forensic treatment just as much as non-cooperating suspects, and avoiding it is their main interest during the course of their legal proceedings. It would be interesting to investigate whether a similar floor effect is found when the DAPP-SF is administered to patients undergoing enforced forensic treatment in high security clinics. This would clarify whether it is indeed the pre-trial nature of the study context that causes the limited ability of the DAPP-SF to screen for personality disorders, or whether it is the forensic setting in general.

As personality disorder plays such a unique role in determining criminal responsibility in the Netherlands, it is important to diagnose it well. A concise screening instrument for personality disorders to successfully identify individuals that require a more thorough and time-consuming assessment could be very helpful in the diagnostic process. However, such an assessment instrument remains to be found, as the DAPP-SF did not succeed in achieving this objective.





Chapter Four

*MMPI profiles of  
males accused of  
severe crimes,  
a cluster analysis*

## Abstract

In studies attempting to classify criminal offenders by cluster analysis of MMPI(-2) data, the number of clusters found varied between ten (the Megargee System) and two (one cluster indicating no psychopathology and one exhibiting serious psychopathology). The latter results raise doubts about the suitability of the MMPI-2 for classification in forensic settings. The present study aimed at deriving an empirical classification system using cluster analysis of 247 MMPI-2 profiles of pretrial criminal defendants in a forensic psychiatric observation clinic. Results indicated only a “non-disturbed” and a “disturbed” profile, differing on general elevation of MMPI-2 profiles but displaying no qualitatively distinct profiles. The clusters differed on age at admission and first conviction, indicating a late onset of criminal activity for disturbed offenders. Also, the clusters differed significantly on Axis I diagnosis and borderline significantly on Axis II diagnosis. The absence of distinct personality profiles between the clusters suggests restricted usefulness of the MMPI-2 in a forensic context of diverse and severe psychopathology and serious crimes. Either the investigated population is in fact extremely homogenous, truly comprising only two kinds of offenders, or the types of offenders in these populations are not effectively distinguished by the MMPI-2.<sup>3</sup>

## Introduction

The Minnesota Multiphasic Personality Inventory-2 (MMPI-2) (Hathaway & McKinley, 1989) is one of the most widely used and researched personality assessment instruments in correctional and forensic psychiatric settings (Nieberding et al., 2003). One line of research in this area has been to attempt to classify criminal offenders into distinct groups according to their MMPI(-2) profiles. Such classification may provide useful information with regard to the motives for committing certain severe crimes, as well as to the treatment and management that is best suited for specific subtypes of offenders.

Among the most influential classification systems is the Megargee System, developed over the past three decades (Megargee et al., 2001). This system, empirically derived in the mid-1970s with the use of cluster analysis on MMPI data of 1,164 subjects, classifies criminal offenders into one of ten clinically informative patterns. Although the system and the classification rules have been tested in a wide variety of criminal justice settings (e.g. Megargee, 1997; Hutton, Miner & Langfeldt, 1993), only a few studies have attempted to replicate the initial cluster analysis leading to this classification system. In fact, whether the MMPI-2 is capable of producing a categorical classification of different types of offenders in other settings remains largely unknown.

Nieberding and colleagues (2003) did, however, replicate Megargee’s findings to a certain extent in a population of 300 mentally disordered inmates preparing for their return to the community. They found seven different clusters, each with distinct personality profiles. Inmates in the first cluster were distinguished by the absence of MMPI-2 clinical scale elevations while the second cluster showed interpersonal and emotional deficits, conflicted relationships, a tendency toward isolation, and a general mistrust of others. The third cluster displayed the highest percentage of Axis II disorders including antisocial features. Inmates in the fourth cluster had the highest percentage of mood disorders such as depression, chronic somatic problems, unresolved anger, and interpersonal conflict. Clinical elevations in the fifth cluster indicated significant personal distress and interpersonal isolation. Nearly half of the inmates in this cluster were diagnosed with schizophrenia, paranoid type. The sixth cluster displayed depressive features such as lethargy, somatic complaints, and anxious features including rumination

.....  
<sup>3</sup> Spaans, M., Barendregt, M., Muller, E., De Beurs, E., Nijman, H., & Rinne, T. (2009). MMPI profiles of males accused of severe crimes: a cluster analysis. *Psychology, Crime & Law*, 15(5), 441-450.

and self-criticism. Individuals in cluster seven were likely to have a history of suicidal and other self-destructive behaviors, episodes of confusion, peculiar or bizarre thoughts, and actively psychotic processes. Consistent with these findings, 88% of inmates in cluster seven received a diagnosis of schizophrenia, which was the highest percentage of all seven clusters (Nieberding et al., 2003).

Other studies, however, have identified far less categories based on MMPI(-2) data. Hall and colleagues (1991) attempted to derive a classification system of sexual offenders in a cluster-analytic procedure using MMPI data and cluster analysis, but found only two clusters: one cluster indicating impulsive and antisocial personality types, and one cluster exhibiting significantly more psychiatric disturbance (Hall et al., 1991).

A similar study aimed at using cluster analysis to identify psychological profiles and related mental health symptoms among juvenile offenders also found just two MMPI clusters (Espelage et al., 2003). The first cluster showed no clinical elevations, indicating the absence of any specific psychological disturbance or distress. The other cluster exhibited profiles of serious psychopathology, with high scores on the scales Schizophrenia, Paranoia, Psychasthenia, and Psychopathic Deviate. Individuals in this cluster were often moody, hostile, and unpredictable, lacked basic social skills, and were prone to violent tempers (Espelage et al., 2003).

Studies indicating only two separate clusters raise doubt about the suitability of using the MMPI-2 for classifying offenders in these settings. Either the investigated populations are extremely homogenous, comprising only two kinds of offenders, or the remaining types of offenders in these populations are not effectively distinguished by the MMPI-2. In the latter case the question arises whether the MMPI-2 should be the primary instrument for differentiating personality types in a forensic population. In the present study, we attempt to derive an empirical classification system based on MMPI-2 profiles in a known heterogeneous population of pretrial criminal defendants in a forensic psychiatric observation clinic. In line with the studies of Megargee and Nieberding, it was hypothesized that in this heterogeneous population, a cluster analysis of MMPI-2 profiles would produce a multitude of distinct personality profiles.

## Method

### **Assessment site**

All subjects were admitted to the Pieter Baan Center (PBC) in Utrecht, the Netherlands, for a residential pre-trial criminal responsibility assessment. The PBC is the official forensic psychiatric observation clinic of the Dutch Ministry of Justice, and has the legal status of a house of detention. The PBC administers around 90% of all clinical forensic assessments in the Netherlands. These assessments cover roughly 5% of all forensic evaluations; the remaining 95% of forensic evaluations take place in a non-specialized forensic setting (usually in a regular house of detention). Possible reasons for the court to order a clinical assessment of defendants include the severity of the crime, the severity of the assumed psychopathology, the maximum security level within the PBC, and potential societal disturbance or media attention associated with the defendant's case. As a result, the population of the PBC covers the more severe criminological and psychiatric cases, and cannot be seen as representative for the entire forensic population whose mental status is assessed.

All defendants are evaluated during a seven-week period by a multidisciplinary team consisting of a psychiatrist, a psychologist, two social workers, and a lawyer who supervises the assessment process along with a second psychiatrist. One of the social workers investigates the life history and social background of the defendant, the other is a supervisor on the defendant's ward whose task is to observe and describe the activities and behavior of defendant during his or her stay in the institution. The psychologist and psychiatrist carry the final responsibility for the PBC's conclusion in the final report concerning DSM-IV diagnoses, if any, and criminal responsibility. The latter two experts also advise the court whether forensic treatment of the defendants is indicated.

### **Subjects**

The sample consisted of 247 defendants, admitted to the PBC between January 1 2000 and December 31 2005 for a criminal responsibility assessment, and of whom an MMPI-2 profile was present. The sample did not include defendants that had refused to participate in the study, and therefore not completed an MMPI-2, and defendants who were unable to complete an MMPI-2 due to intellectual disabilities, a poor comprehension of the Dutch language, or the presence of psychotic symptoms. Because of the low number of female defendants admitted to the institution (less than 10%), it was decided to include only male subjects in the study.

### **Measures**

The variables comprise three categories of information: individual characteristics, clinical characteristics, and MMPI-2 data.

#### **Individual characteristics**

Individual characteristics included demographic variables (e.g., age of defendant at time of assessment, number of earlier convictions, age at first conviction, whether the defendant was a first offender, and cultural background of the defendant), and crime related variables (e.g., type of index crime, relationship between the perpetrator and the victim [i.e., acquainted or not acquainted with the victim], type of weapon used, information from the crime scene, and whether the offence was committed with others or alone).

For first offenders, the age at the time of the assessment was taken as a proxy for age at first conviction. For all defendants charged with homicide, the variable "attempted murder" indicated whether the charge involved *attempted* homicide (i.e., not lethal to the victim) or actual homicide (i.e., lethal to the victim).

#### **Clinical characteristics**

Clinical characteristics included the presence or absence of DSM-IV Axis I psychiatric diagnosis and Axis II personality disorder. Axis I substance abuse was scored on the basis of the official diagnosis in the final report. In Dutch forensic practice, however, substance abuse is assessed only when judged clinically relevant for the index crime. As a result, there may be some underreporting compared to actual substance abuse (e.g., nicotine abuse).

### **Minnesota Multiphasic Personality Inventory-2**

The MMPI-2 (Hathaway & McKinley, 1989), is a self-report questionnaire consisting of 567 True or False questions designed to detect psychopathologic symptoms in psychiatric patients. Included in this study were T-scores on three validity scales as well as the basic clinical scales. The Validity scales are: L (Lie, unsophisticated lying), F (Low Frequency, tendency to answer affirmative to items rarely endorsed by normal people, also indicating psychopathology), and K (Correction, sophisticated lying). Clinical scales include Hs (Hypochondrias), D (Depression), Hy (Hysteria), Pd (Psychopathic deviation), Mf (Masculinity-Femininity, measuring the subjects identification with traditional gender roles), Pa (Paranoia), Pt (Psychastenia, indicating inter alia compulsions, obsessions, abnormal fears and difficulties in concentration), Sc (Schizophrenia), Ma (Hypomania), and Si (Social Introversion).

### **Statistical analyses**

An inter-rater reliability analysis was carried out on twenty files to determine the degree of consensus on the scoring of individual and clinical characteristics. Total agreement was found for Axis I substance abuse, cultural background of the defendant, whether the defendant was a first offender, degree of criminal responsibility, and presence of an Axis I and/or II disorder (Cohen's Kappa = 1.00). Other variables demonstrated satisfactory agreement; type of weapon used (K = .92), location of the crime (.63), presence of a personality disorder (.57), age of the victim (K = .77), presence of an Axis I disorder (K = .90), and type of crime (K = .83).

A hierarchical cluster analysis was performed using Ward's method (squared Euclidean distance). Hierarchical cluster analysis is a statistical method aimed at creating objective classifications within a dataset, grouping together subjects with similar patterns on clustering variables. As we were interested in clinically identifiable groups, the MMPI-2's clinical scales were chosen as clustering variables. Ward's method (Ward, 1963) of cluster analysis initially treats all subjects as separate clusters. In the second step, the two subjects that are most identical on the clustering variables are clustered. To determine the two most identical subjects, the within cluster error sum of squares (ESS) of every possible combination is calculated and the result with the least ESS is chosen. This procedure is repeated by either clustering two subjects into a new cluster, a subject into an already existing cluster, or clustering two already existing clusters into a new cluster, until in the last step all subjects are combined into one all-encompassing cluster.

When clusters are merged, the within cluster variance will naturally become larger and the ESS will increase. When two clusters that are much alike are joined, the increase in ESS will be relatively small. When two clusters that are very different are joined, the increase will be relatively large. By carefully observing the way the ESS increases in each step of the cluster analysis, one can determine the appropriate number of clusters – the step before a sudden jump in ESS being the best point to terminate further clustering. As the criteria to determine the appropriate number of clusters are always based on professional judgment and therefore remain somewhat subjective, a number of alternative cluster solutions were also investigated to establish whether these resulted in clinically interpretable clusters.

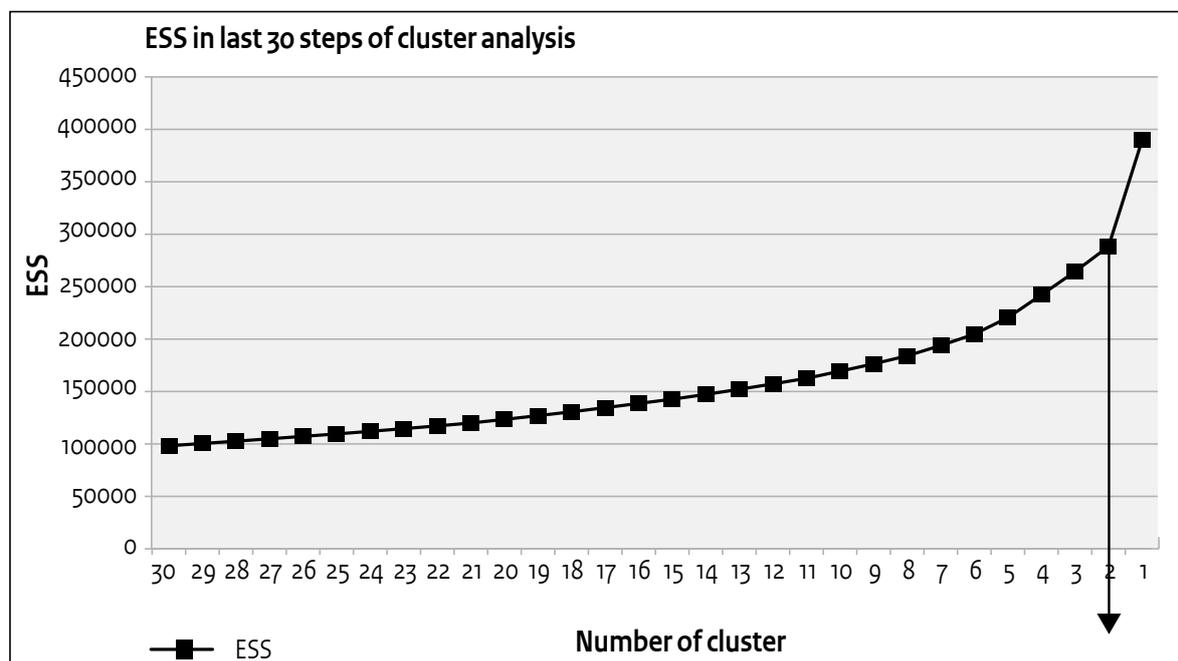
Hierarchical clustering is useful in determining the best fitting number of clusters, but may perform wrongly when assigning individual cases to clusters due to the agglomerative nature of the process (Huberty, DiStefano, & Kamphaus, 1997). With every clustering, the centroid of the cluster changes, as a result of which cases that have been assigned to a certain cluster in the beginning of the process might better be assigned to another cluster at the end of the process. In order to assign cases to clusters, a k-means cluster analysis is used, which iteratively allocates all cases to a specified (*k*) number of clusters based on each cluster's nearest center points.

To investigate differences between clusters on clinical, demographic, and crime related characteristics, logistic regression and chi-square analyses were used.

## Results

Figure 1 shows the agglomeration schedule of the hierarchical cluster analysis for the last 30 steps of the procedure. A sudden increase in ESS appeared at the final step, indicating that a two cluster solution would be the most appropriate. The first cluster showed considerably higher scores on all MMPI-2 subscales than the second cluster, but no distinguishable personality profiles were found.

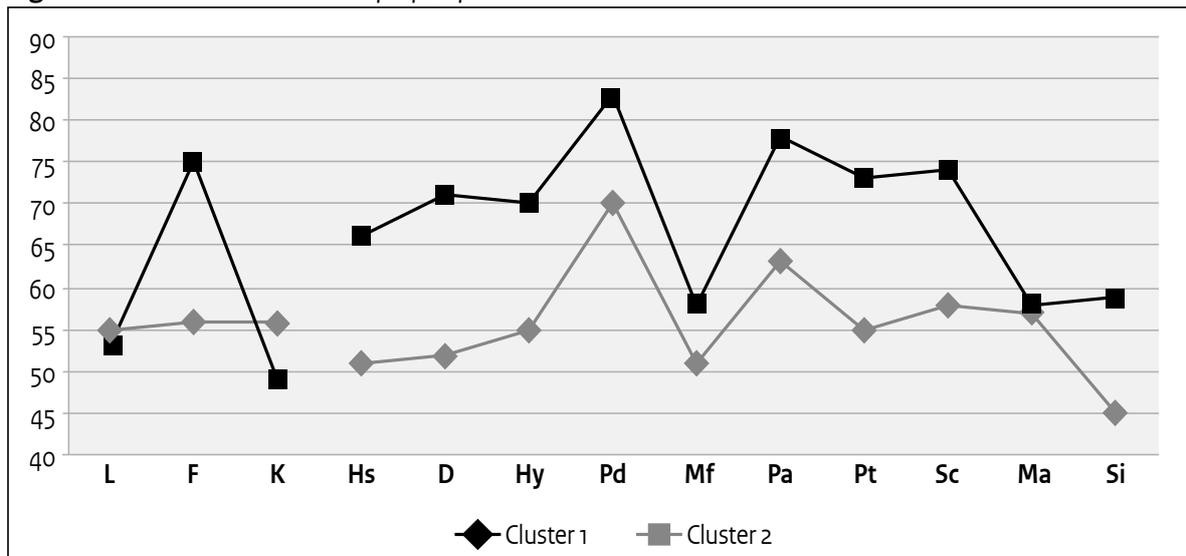
**Figure 1.** Agglomeration schedule in last 30 steps of cluster analysis



The total within clusters error sum of squares (ESS) of the last 30 steps of the hierarchical cluster analysis. The sudden increase in ESS in the last step suggests that the two cluster solution is the appropriate one.

As can be seen in Table 1, the two cluster solution comprises one cluster with only a clinical elevation on scale 4 (Cluster 1, “non-disturbed”) and one cluster with elevations on all scales except scales 5, 9, and 0 (Cluster 2, “disturbed”).

**Figure 2.** Mean MMPI-2 cluster profiles for 2 cluster solution



Furthermore, the clusters’ profiles appear surprisingly similar with the three highest elevations on scales 4, 6, and 8 in both clusters (Figure 2).

**Table 1.** Clusters' mean MMPI-2 profiles for four cluster solutions

	<i>n</i> (%)	L	F	K	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma	Si	Mean code
<b>2 clusters</b>															
Cluster 1	156	55	56	56	51	52	55	70	51	63	55	58	57	45	4-6
Cluster 2	91	53	75	49	66	71	70	83	58	78	73	74	58	59	4-6
<b>3 clusters</b>															
Cluster 1	86	56	51	57	48	50	52	64	50	61	53	54	53	44	4-6
Cluster 2	52	51	81	47	70	77	75	86	58	81	78	78	57	64	4-6
Cluster 3	109	54	63	54	57	58	60	79	55	69	61	65	61	48	4-6
<b>4 clusters</b>															
Cluster 1	77	54	69	55	61	66	66	81	58	75	69	70	57	56	4-6
Cluster 2	99	55	56	59	54	53	59	75	51	63	58	60	59	42	4-6
Cluster 3	50	55	54	51	45	48	47	59	49	63	50	54	53	48	4-6
Cluster 4	21	50	94	41	76	86	80	88	60	88	85	87	62	43	4-6
<b>5 clusters</b>															
Cluster 1	5	56	53	43	36	41	44	52	43	66	30	38	54	40	6-9
Cluster 2	53	51	62	52	51	48	51	74	53	63	56	62	72	41	4-6
Cluster 3	74	55	63	56	62	64	68	82	57	73	67	67	55	52	4-6
Cluster 4	85	57	54	58	51	54	56	67	49	62	56	57	49	47	4-6
Cluster 5	30	52	93	40	72	82	75	85	60	85	82	84	62	69	4-6

Validity scales: L (Lie), F (Low Frequency), and K (Correction). Clinical scales: Hs (Hypochondrias), D (Depression), Hy (Hysteria), Pd (Psychopathic deviation), Mf (Masculinity-Femininity), Pa (Paranoia), Pt (Psychastenia), Sc (Schizophrenia), Ma (Hypomania), and Si (Social Introversion). The clustering variables were the clinical scales 1 to 0. The validity scales L, F & K were not part of the clustering procedure, but are presented here as descriptive information. The cluster's mean code is presented as descriptive information only and is based on the rules for determining the high-point-code (McGrath, Rashid, Hayman, & Pogge, 2002): 1) the code was based on the two most elevated clinical scales, 2) scales 5 and 0 were ignored, 3) numeric precedence was used to resolve ties, and 4) the order of elevation was ignored.

Three alternative solutions were investigated in order to ascertain whether these would yield different profile structures, which they did not. For all alternative solutions, the main between-cluster difference seemed to be the general elevation of each profile, not a qualitative difference between the profiles themselves. Only in the five-cluster solution did one cluster emerge with a somewhat different profile, although it must be mentioned that this was a small cluster of only five subjects.

Results of the regression analysis are shown in Table 2. Initial tests for colinearity revealed a very high correlation between the variables defendant's age at first conviction and age of the defendant at time of assessment ( $r = .681$ ). Because of this high correlation and the conceptual overlap between the two variables it was decided to include only the variable age at first conviction. The two clusters found in the initial cluster analysis did not differ on most clinical, offender and offence characteristics. Clinically, the two clusters only differed significantly on whether the subjects received an Axis I diagnosis and borderline significantly on receiving an Axis II diagnosis. Furthermore, age at first conviction was significantly higher in the disturbed cluster than in the non-disturbed cluster.

**Table 2.** Logistic regression summary for diagnostic, offender, and offence characteristics

Variable	$\beta$	SE	p	OR
Received Axis I diagnosis	.754	.335	.024	2.125
Received Axis II diagnosis	.614	.368	.095	1.849
Received Axis II diagnosis substance abuse	.242	.335	.471	1.273
First Offender	-.348	.431	.420	0.706
Number of prior convictions	.035	.032	.273	1.036
Age at first conviction	.058	.020	.003	1.060
Cultural minority	.398	.324	.220	1.489
Type of offence (1)				
Homicide	.093	.442	.834	1.097
Sexual offence	.522	.547	.340	1.686
Offence at offender's home	-.555	.378	.142	0.574
Offended with others	-.294	.443	.506	0.745
Victim known to offender	-.329	.361	.362	0.720
Victim's age < 16	-.386	.474	.415	0.680
Weapon choice (2)				
Knife	.092	.438	.834	1.096
Firearms	.176	.527	.739	1.192
Other	.483	.456	.290	1.620

(1) Reference category: other. (2) Reference category: none (only physical)

The two clusters did not differ significantly in level of criminal responsibility (see Table 3). However, the disturbed group did receive the advice of forensic treatment from the PBC's experts more often than the non-disturbed group.

**Table 3.** *Intercluster differences for expert's advice*

Expert's advice	Cluster 1	Cluster 2	$\chi^2$ (df)	P
Criminal responsibility			5.68 (4)	.225
Responsible	19 (12%)	8 (9%)		
Slightly diminished responsible	39 (25%)	13 (14%)		
Diminished responsible	77 (49%)	56 (62%)		
Strongly diminished responsible	15 (10%)	9 (10%)		
Not responsible	6 (4%)	5 (6%)		
Measure advised	82 (53%)	66 (73%)	9.54 (1)	.002

## Discussion

Based on MMPI-2 information, only two groups of distinct personality profiles were found: a “non-disturbed” and a “disturbed” profile. This conclusion supports the findings by Espelage et al. (2003) of one cluster indicating the absence of any psychological problems and the other exhibiting serious psychopathology. Even in a diverse population – containing not only severe and varied psychopathology but also individuals with no psychological disorders – different types of offenders are simply not effectively distinguished by the MMPI-2. Furthermore, the two resulting clusters did not reveal different types of offenders in the sense of different MMPI-2 profiles, but only varied in the degree of general psychopathology. These findings raise serious doubts about the usefulness of interpreting MMPI-2 profiles for differential diagnostic purposes. It may be concluded that the usefulness of the MMPI-2 in a forensic context is restricted to screening for the presence or absence of general psychopathological symptoms. Of particular interest, and reason for further caution in interpreting MMPI-2 profiles, is the fact that the two clusters only differed significantly on Axis I symptoms while the results for Axis II personality disorders were only borderline significant.

Two notable intercluster differences are worth mentioning, however. Firstly, the subjects in the disturbed cluster started their criminal careers later in life, indicated by the age at first conviction, than the non-disturbed cluster. This outcome concurs with findings of a study by Nijman, Cima and Merkelbach (2003) that the age at first conviction of psychotic offenders was significantly higher than that of non-psychotic offenders (i.e. psychotic offenders of severe crimes were first offenders more often than non-psychotic offenders), indicating that the psychotic offenders had a late onset of criminal activity. Furthermore, according to Nijman and colleagues (2003), a large portion of the psychotic or disturbed offenders had received psychiatric treatment prior to their first offense, indicating that they had a substantial psychiatric history, while the non-disturbed offenders typically came from a long criminal career and can be described as early starters. In line with these findings, it might also be the case that disturbed offenders are less exposed to opportunities to commit crimes at a young age because they are often under protective orders until at least the age of 18.

Secondly, the disturbed group received an advice for forensic treatment more often than the non-disturbed group. This finding itself is not surprising – the aim of such a measure is to treat the mental disorder that influenced the alleged crime, and more importantly, a mental disorder is one of the main criteria for advising forensic treatment. It confirms that the MMPI-2 is effective at measuring whether overall psychopathology was elevated in the present population. However, the fact that no profiles are retrievable empirically indicates that, in this particular population, the MMPI-2 is not able to distinguish more complex psychopathology.

A limitation of the study is the inclusion of only males in the study population. Also, not all willing observees are able to complete an MMPI-2 due to limitations such as intellectual disabilities or psychotic symptoms. Perhaps the inclusion of MMPI-2 profiles of these observees could have resulted in more than two clusters. However, these profiles were never realized as the MMPI-2 is not well-suited for individuals with the above-mentioned problems. Furthermore, as mentioned earlier, the clinical assessments performed in the PBC are only a small and probably highly specific proportion of all forensic assessments conducted in the Netherlands. Therefore, the results of this study can only be generalized to defendants of very severe crimes.

Despite the cluster analysis of 247 MMPI-2 scores not resulting in distinct personality type profiles, the MMPI-2 can certainly be used in individual cases to investigate the presence of psychopathology. Nonetheless, implications for future research in a similar population include a detailed investigation of individual MMPI-2 profile results and of the NEO-PI-R, a possible alternative instrument that can differentiate personality types in a forensic population.



# 5

Chapter Five

*Diagnosis of anti-social personality disorder and criminal responsibility*

## Abstract

The present study empirically investigates whether personality disorders and psychopathic traits in criminal suspects are reason for diminished criminal responsibility or enforced treatment in high security hospitals. Recently, the tenability of the claim that individuals with personality disorders and psychopathy can be held fully responsible for crimes has been questioned on theoretical bases. According to some interpretations, these disorders are due to cognitive, biological and developmental deficits that diminish the individual's accountability. The current article presents two studies among suspects of serious crimes under forensic evaluation in a Dutch forensic psychiatric observation clinic. The first study examined how experts weigh personality disorders in their conclusions as far as the degree of criminal responsibility and the need for enforced forensic psychiatric treatment are concerned ( $n = 843$ ). The second study investigated associations between PCL-R scores and experts' responsibility and treatment advisements ( $n = 108$ ). The results suggest that in Dutch forensic practice, the presence of a personality disorder decreased responsibility and led to an advice for enforced forensic treatment. Experts also take characteristics of psychopathy concerning impulsivity and (ir)responsibility into consideration when judging criminal accountability. Furthermore, they deem affective deficiencies sufficiently important to indicate suspects' threat to society or dangerousness and warrant a need for forensic treatment.<sup>4</sup>

## Introduction

Full criminal responsibility implies that an individual who commits a crime was fully aware of the (illegal) nature, character and consequences of that crime. When an individual suffers from a severe mental disorder that leads to a crime, it is generally agreed in most jurisdictions that he or she cannot be held criminally responsible for it and should be exempt from its penal consequences. A number of countries, such as Canada and a number of U.S. states, use a dichotomy of options when it comes to criminal responsibility. An offender is viewed either as fully responsible and receives a prison sentence, or the crime was the result of a mental disorder, the offender is viewed as criminally insane, and the court imposes enforced treatment in a high security forensic psychiatric hospital. Elsewhere, a graded system is used, allowing for various possible grades of criminal responsibility.

Not all mental disorders are considered a potential cause for diminished criminal responsibility. In many jurisdictions the mere presence of a personality disorder is not viewed as sufficient grounds for criminal insanity and forensic treatment. This holds especially for antisocial personality disorder and psychopathy with crucial diagnostic criteria such as criminal versatility and repeated unlawful behaviors. Some authors have questioned, however, the tenability of the claim that individuals with personality disorders can be held fully responsible. They argue that personality disorders and psychopathy *can* be interpreted as serious mental disorders, based as they are on developmental disabilities or particular deficits such as cognitive deficiencies and biological impediments. Mei-Tal (2002), for instance, argued that the complete absence of empathy in persons with high psychopathy scores implies that they should never be regarded as responsible agents or blameworthy. Earlier, Herpertz and Sass (2000) concluded that violence in persons with high psychopathy scores is rooted in emotional deficiency. Due to deficient emotional learning they show poor conditioning processes, cannot be conditioned to avoid punishment, and are unable to evaluate the consequences their actions will have. They therefore make no effort to avoid harmful behavior or suppress violent

.....  
<sup>4</sup> Spaans, M., Barendregt, M., Haan, B., Nijman, H., & De Beurs, E. (2011). Diagnosis of antisocial personality disorder and criminal responsibility. *International Journal of Law and Psychiatry*, 34, 374-378.

impulses. Emotional deficiency is also associated with a general under-arousal, which in turn may lead to sensation seeking and risk-taking in the form of violence or other illegal behavior (Herpertz & Sass, 2000). According to Ciocchetti (2003), punishment is inappropriate for persons with high psychopathy scores due to their failure to understand the significance and influence of their responses to the acts of others. They cannot appropriately interpret punishment because they cannot understand the wrongfulness of their actions or the significance of any punishment they are given. Fine and Kennett (2004) argued that psychopathic offenders are incapable of forming genuine moral concepts because they failed to pass through a crucial moral developmental stage in early childhood and therefore cannot meet the requirement of being criminally responsible. Palermo (2007) argued that under conditions of severe stress, individuals suffering from psychopathic or antisocial personality disorders may decompensate and experience either fleeting or short-term psychotic thinking and behavior that can severely impair the ability to reason or act rationally, to distinguish right from wrong, and to conform to the law. According to Palermo, the best legal option in such a case would be commitment to a mental forensic institution for suitable treatment.

### **Criminal responsibility in the Netherlands**

Under current Dutch criminal law, a crime committed due to a disorder that rendered the offender unable to act differently and the offense unavoidable is not considered punishable. According to Barendregt (2006), this legal decision finds its origin in the traditional image of man, which includes the idea that individuals can, to a certain degree, take control over their actions. Without any such degree of freedom of will, individuals cannot be held responsible for their actions, either personally or criminally. Freedom of will exists in the ability of humans to self-reflect. Reflective understanding, or self-consciousness, allows for actions to be planned, controlled, inhibited, reviewed and revised. Hence, the ability to self-reflect indicates the freedom to perform an act or to refrain from it. Should this capacity have been affected by a mental disorder, an individual's actions are not completely planned or controlled. As a consequence, responsibility for that action may be diminished. The court may take the diminished responsibility into account and, for instance, decide for a mitigated sentence or compulsory treatment.

Dutch criminal practice distinguishes five possible degrees of criminal responsibility, as opposed to the sane-insane dichotomy used in some other jurisdictions. The 5-point scale ranges from complete responsibility, slightly diminished, diminished, severely diminished to total absence of responsibility. These grades of responsibility are related to the intensity of the role played by a psychiatric disorder, if found, in the crime. The more serious the disorder and the role it played in the crime, the less responsible an offender will be held. The Dutch dimensional approach to criminal responsibility fits in well with the dimensional nature of psychiatric or personality disorders. Mental illness is not an all-or-nothing phenomenon, but symptoms wax and wane, sometimes reaching levels above disorder threshold (so-called syndromal states). This is especially the case in personality disorders which has led to the proposal of a dimensional model of classification (Widiger, 2000). Dimensional diagnoses reflect the differences in the number and severity of disorder criteria with the diagnostic groups of persons below the disorder threshold (Ullrich, Borkenau & Marneros, 2001). However, the fine grained categorization of criminal responsibility has also been criticized, as it suggests an accuracy and reliability of its assessment that cannot be guaranteed in practice. Nevertheless, it is firmly founded in Dutch forensic practice and in the judiciary system. In cases of diminished or severely diminished

criminal responsibility, courts in the Netherlands can impose an equivalently diminished prison sentence followed by enforced treatment in a high security forensic psychiatric hospital.

Prior to the trial the court can request a forensic psychiatric evaluation. Around 90% of all *inpatient* forensic assessments in the Netherlands are carried out by the Pieter Baan Center (PBC), the official forensic psychiatric observation clinic of the Dutch Ministry of Justice. These assessments cover roughly 5% of all forensic evaluations; the remaining 95% are regular *outpatient* evaluations that take place in a non-specialized forensic setting (usually where the defendant is being detained). Possible reasons for the court to order such a specialized inpatient assessment of defendants in the PBC include the severity of the crime, the severity of the assumed psychopathology, the maximum security level within the PBC, and potential societal disturbance or media attention associated with the defendant's case. As a result, the population of the PBC covers the more severe cases as far as criminological and psychiatric backgrounds are concerned. All defendants are evaluated during a seven-week period by a multidisciplinary team consisting of a psychiatrist, a psychologist, two social workers, and a lawyer who supervises the assessment process along with a second psychiatrist. One of the social workers investigates the life history and social background of the defendant through interviews with informants such as family members, the other is a supervisor on the defendant's ward whose task is to observe and describe the activities and behavior of the defendant during his or her stay in the institution. The psychologist and psychiatrist carry the final responsibility for the PBC's conclusion in its report concerning DSM-IV psychiatric diagnoses, if any, and criminal responsibility (based on structured instruments and clinical judgment). The latter two experts also advise the court whether forensic treatment of the defendant is indicated if convicted of the charge.

Existing research from the U.S. and Canada demonstrated that (antisocial) personality disorder decreased the chance that a defendant was judged insane, which is in line with current legislation in these countries (Warren et al., 2004; Rice & Harris, 1990). However, Barendregt, Muller, and colleagues (2008) found that in the expert opinion of Dutch forensic psychologists and psychiatrists, a personality disorder, while decreasing the chances of total absence of responsibility, was associated with a judgment of diminished responsibility as well as a higher chance of enforced forensic treatment in a high security hospital. These results suggested that the sane-insane dichotomy used in many jurisdictions around the world may be in need of revision. A more nuanced categorization, might better allow for the subtle role that psychopathy or personality disorders can play in matters such as criminal responsibility and freedom of will.

### **Research questions**

This study aims to examine how, in pre-trial psychiatric forensic evaluations, Dutch experts weigh the presence of a personality disorder in their conclusions regarding criminal responsibility compared to other psychiatric conditions, and how they advise on the necessity of enforced treatment in a high security hospital for personality disorders. Furthermore, how do Dutch experts weigh various aspects of psychopathy as measured by the Psychopathy Checklist-Revised (PCL-R), in the above-mentioned judgment regarding criminal responsibility and how do they advise regarding the need for enforced treatment? To answer these questions, two studies using the files of the Pieter Baan Center were carried out.

# *STUDY 1: How do experts weigh personality disorder in their forensic evaluations?*

## Method

A total of 1209 reports of defendants, admitted to the PBC between January 1, 2002 and December 31, 2007, were examined to obtain information on the presence of psychiatric disorders, the conclusion regarding criminal responsibility and the experts' advice on possible forensic treatment in a high security hospital. Forty files could not be retrieved from the PBC's archive at the time of the study and in 319 reports the team of forensic experts was not able to carry out a full forensic evaluation due to the defendants' refusal to cooperate, which left 850 reports. After listwise removal of missing values a total of 843 complete cases remained for Study 1. The sample consisted of 89% men and 11% women with a mean age of 33.1 years ( $SD = 10.2$ ). Ethnicity was primarily Dutch (53.3%), followed by Surinamese (10.9%), Moroccan (6.2%), Dutch Antillean (5.8%), and Turkish (5.3%). The sample also consisted of defendants of other (non-Dutch) European origin (2.1%) and those of other (non-European) origins than those mentioned above (16.5%).

### **Variables**

The dependent variables were expert's opinion regarding (1) criminal responsibility according to the aforementioned 5-point scale and (2) their advice for enforced treatment in a high security psychiatric hospital as a dichotomous variable (yes/no). The five possible variations on the categorical scale of criminal responsibility are complete responsibility, slightly diminished responsibility, diminished responsibility, severely diminished responsibility, and total absence of responsibility. Independent variables were presence of a psychotic disorder, presence of a personality disorder, presence of a substance abuse disorder, co-morbidity of psychotic and personality disorder, and an IQ below 80. All independent variables were dichotomous (yes/no).

### **Statistical analyses**

Chi-square tests were carried out to determine whether there were significant differences between the responsibility categories and the presence of a psychotic disorder, personality disorder, substance abuse disorder, co-morbidity of psychotic and personality disorder, and an IQ below 80. Also, the association between advice for enforced treatment and the presence of these disorders was tested with a chi-square test. A multinomial logistic regression was carried out to verify whether the presence of these disorders predicts the degree of criminal responsibility (a categorical variable with five levels). A binary logistic regression was carried out using a forced entry model to determine the relationship between the presence of the disorders and advice for enforced treatment in a high security hospital, using the advice for enforced treatment as a dependent binary variable. P-values below .05 are reported as statistically significant.

# Results

## Descriptive statistics

Statistically significant differences between the responsibility categories were found for the presence of a psychotic disorder, substance use disorder, personality disorder, and co-morbidity of these disorders, as shown in Table 1. For the 150 subjects who were deemed fully responsible for their indicted crime, 51 had a personality disorder (and 99 had no personality disorder). For the 70 subjects who were deemed not responsible for their indicted crime, 9 had a personality disorder (and 61 did not). Inspection of the remaining frequencies showed that the presence of a psychotic disorder was primarily associated with severely diminished criminal responsibility ( $n = 62$ , 55%) and full absence of responsibility ( $n = 66$ , 94%). In contrast, the presence of a personality disorder was especially frequent in the category of diminished responsibility ( $n = 289$ , 79%) and slightly diminished ( $n = 87$ , 61%).

**Table 1.** Chi-square tests for diagnostic variables compared to degrees of criminal responsibility

	Criminal responsibility					$\chi^2(1)$	P
	FR	Sl. Dim	Dim	Sev. Dim.	NR		
n	150	141	369	113	70		
Personality disorder	51	87	289	46	9	170.13	< .000***
Psychotic disorder	10	11	50	62	66	319.24	< .000***
Substance abuse	41	61	187	54	33	24.37	< .000***
Co-morbidity	2	5	24	19	7	26.64	< .000***
IQ < 80	4	19	51	16	9	14.59	.006**

FR = Full responsibility; Sl. Dim = Slightly diminished responsible; Dim = Diminished responsible; Sev. Dim = Severely diminished responsible; NR = Not responsible. (1)  $df = 4$ ; \*\* =  $p < .01$ ; \*\*\* =  $p < .001$

A multinomial logistic regression was conducted using diagnostic variables as predictors for the responsibility category (see Table 2). The presence of a personality disorder predicted slightly diminished, diminished and severely diminished responsibility but not complete absence of responsibility

(as compared to full responsibility). Inspection of the effect-sizes demonstrates that the presence of a personality disorder was most important for the middle of the spectrum (diminished responsibility) gradually losing impact towards both extremes. The presence of a psychotic disorder predicted all categories except slightly diminished responsibility when compared to full responsibility. Like poor intellectual functioning (IQ < 80), the presence of a psychotic disorder was a straightforward predictor for criminal responsibility with increasing effect-sizes towards the not responsible endpoint. The presence of a substance use disorder differentiated only for the diminished responsibility category.

**Table 2.** Multinomial Logistic Regression Analysis of 843 Criminal Responsibility Opinions

Dependent category Predictor	b	SE b	Wald's $\chi^2$ (df=1)	p	OR
<i>Slightly diminished responsible</i>					
Constant	-.960	.201			
Psychotic disorder	.435	.573	.575	ns	1.545
Substance use disorder	.476	.261	3.331	ns	1.610
Personality Disorder	1.182	.263	20.264	< .000	3.261
Co-morbidity	-.107	1.031	.011	ns	.899
IQ < 80	1.985	.573	12.016	.001	7.267
<i>Diminished responsible</i>					
Constant	-.935	.194			
Psychotic disorder	1.803	.446	16.351	< .000	6.066
Substance use disorder	.603	.230	6.869	.009	1.828
Personality Disorder	2.239	.242	85.659	< .000	9.368
Co-morbidity	-1.096	.877	1.564	ns	.334
IQ < 80	2.217	.551	16.194	.000	9.180
<i>Severely diminished responsible</i>					
Constant	-1.756	.260			
Psychotic disorder	3.139	.458	46.971	< .000	23.081
Substance use disorder	.563	.292	3.719	ns	1.757
Personality Disorder	.786	.343	5.255	.022	2.195
Co-morbidity	-3.79	.907	.175	ns	.685
IQ < 80	2.304	.598	14.843	< .000	10.012
<i>Not responsible</i>					
Constant	-4.354	.746			
Psychotic disorder	6.035	.826	53.443	< .000	417.806
Substance use disorder	.522	.370	1.984	ns	2.1.685
Personality Disorder	.740	1.024	.521	ns	2.095
Co-morbidity	-1.673	1.360	1.514	ns	.188
IQ < 80	2.710	.715	14.348	< .000	15.025

Note. Reference Category = Fully responsible. Model:  $\chi^2(20) = 459.374, p < .000$ . Deviance Goodness of Fit:  $\chi^2(40) = 42.581, p = .361$ . Cox and Snell  $R^2 = .420$ . Nagelkerke  $R^2 = .445$ .

### Need for enforced treatment

As shown in Table 3, enforced treatment in a high security hospital was advised in slightly over half of all cases. All diagnostic variables but intellectual functioning differed significantly between the groups that did and did not receive such an advice. Inspection of the frequencies showed that the presence of a personality disorder was associated with an advice for enforced treatment. The same held true for the presence of a psychotic disorder, the presence of a substance use disorder, and co-morbidity of these disorders.

**Table 3.** Chi-square tests for diagnostic variables compared to advice for enforced treatment

ccc	Enforced treatment			
	No	Yes	$\chi^{2(2)}$	P
n	433	410		
Personality disorder	228	254	7.43	.006**
Psychotic disorder	51	148	69.07	< .000***
Substance abuse	164	212	16.31	< .000***
Co-morbidity	15	42	15.35	< .000***
IQ < 80	44	55	2.15	.143

(2)  $df = 1$ ; \*\* =  $p < .01$ ; \*\*\* =  $p < .001$

A binary logistic regression analysis was conducted using the advice for enforced treatment as dependent variable. The results in Table 4 show that when controlled for the other variables, the presence of a personality disorder was also positively related to an advice for enforced treatment in a high security hospital.

**Table 4.** Logistic Regression Analysis of 843 assessments for expert's advice of enforced intramural treatment

Predictor	b	SE b	Wald's $\chi^2$ (df = 1)	p	OR
Constant	-1,406	.174	65.383		
Psychotic disorder	2.297	.255	81.229	< .000	9.943
Substance use disorder	.344	.152	5.137	.023	1.411
Personality Disorder	1.173	.193	36.952	< .000	3.231
Co-morbidity	-1.310	.407	10.346	.001	.270
IQ < 80	.512	.236	4.720	.030	1.669

Note. Reference Category = no advice for enforced treatment. Model:  $\chi^2(4) = 107.504$ ,  $p < .000$ . Deviance Goodness of Fit: Chi-square: 1.471(3),  $p = .689$ . Cox and Snell  $R^2 = .166$ . Nagelkerke  $R^2 = .221$ .

## STUDY 2: *How do experts weigh psychopathy in their forensic evaluations?*

### Method

To answer the question how experts weigh various aspects of a psychopathic personality in their evaluations, the Psychopathy Checklist-Revised (PCL-R) was administered on file information of a subsample of the sample in Study I. Criminal responsibility assessments reported within one selected year (from September 1, 2006 to August 31, 2007) were selected, excluding reports on female suspects as the PCL-R is not well suited for women. This resulted in a sample of 108 reports on male suspects with a mean age of 33.2 years ( $SD = 10.6$ ). Ethnicity was primarily Dutch (57.1%), followed by Surinamese (10.2%), Moroccan (7.1%), Dutch Antillean (1.0%), Turkish (5.1%), other European (1.0%), and other non-European (18.3%).

The PCL-R was administered by two trained individuals based on the report as well as other file information (such as detailed description of the crime provided by the police and criminal investigators, (psycho)medical correspondence, and psychological test results). The files provided sufficient information for reliable administration of the PCL-R; intra-class correlation coefficient (ICC) based on 50 assessments was .85 for PCL-R total score, .72 for Factor 1, .82 for Factor 2, .67 for Facet 1, .62 for Facet 2, .74 for Facet 3, and .87 for Facet 4 scores.

#### **The PCL-R**

The PCL-R consists of twenty items which are each allocated a score of 0-2 making the highest possible total score 40. In Europe, psychopathy is considered present at a score of 26 or higher (Grann, Långström, Tengström & Stålenheim, 1998). Seventeen of the twenty items are divided over two factors. The PCL-R was scored according to Hare's two-factor/four-facet model. Factor one is characterized by selfish, callous, and remorseless use of others (an egotistical interpersonal style and shallow affective features), and factor two measures a chronically unstable and antisocial lifestyle. Facet 1 comprises an arrogant and deceitful interpersonal style, facet 2 deficient affective experience, facet 3 an impulsive and irresponsible behavior style, and facet 4 antisocial behavior (Hare, 2006).

#### **Variables**

Expert's opinion on criminal responsibility was operationalized according to the aforementioned 5-point scale (see Study 1). Expert advice for enforced treatment in a high security psychiatric hospital was a dichotomous variable (yes/no). PCL-R total, factor, and facet scores were collected based on file information from PBC assessments.

#### **Statistical analyses**

Statistical analyses included a one-way ANOVA to compare the PCL-R item, facet, and total scores to the five criminal responsibility categories. A t-test was carried out on PCL-R item, facet, and total scores to ascertain statistically significant differences between reports with and without a positive advice for enforced treatment. P-values below .05 are reported as statistically significant.

## Results

### Criminal responsibility

No significant differences were found between the five levels of criminal responsibility for the facet or total score on the PCL-R (see Table 5). At an item-level, high scores on three PCL-R items yielded statistically significant results: impulsivity ( $F(4,103) = 5.287, p = .001$ ), irresponsibility ( $F(4,106) = 2.517, p = .046$ ), and failure to accept responsibility for one's own actions ( $F(4,107) = 2.723, p = .033$ ) (data not presented in the table; results for remaining items available on request).

**Table 5.** ANOVA for 108 PCL-R facet and total scores

Criminal responsibility							
	FR	Sl. Dim	Dim	Sev. Dim.	NR	F	p
n	20	20	41	14	13		
Facet 1	2.23 (2.2)	2.30 (1.6)	2.38 (2.1)	2.36 (1.6)	1.85 (1.7)	.198	.939
Facet 2	4.60 (2.7)	5.03 (2.0)	5.84 (1.8)	5.55 (1.9)	4.62 (1.9)	1.740	.147
Facet 3	3.93 (2.8)	4.31 (2.0)	4.87 (2.4)	5.89 (2.1)	5.71 (3.0)	1.977	.104
Facet 4	2.40 (2.7)	3.20 (2.9)	3.65 (2.8)	3.79 (2.0)	3.04 (2.0)	.917	.457
Total	13.98 (8.4)	15.96 (7.3)	17.74 (7.6)	18.55 (5.8)	15.63 (6.4)	1.210	.311

Mean scores (SD) for PCL-R Facets and Total. FR = Full responsibility; Sl. Dim = Slightly diminished responsible; Dim = Diminished responsible; Sev. Dim = Severely diminished responsible; NR = Not responsible.

### Need for enforced treatment

The results of the t-test, carried out on PCL-R item, facet, and total scores to ascertain statistically significant differences in the advice for enforced treatment, show that high scores on Facet 2 predicted an advice for enforced treatment ( $t = -2.523, p = .013$ ), as did high scores on Facet 3 ( $t = -2.361, p = .020$ ) and a high total score ( $t = -2.293, p = .024$ ) (see Table 6). High scores on a number of items also yielded statistically significant differences: lack of remorse or guilt ( $t = -2.382, p = .019$ ), callousness and lack of empathy ( $t = -2.265, p = .026$ ), poor behavioral controls ( $t = -2.961, p = .004$ ), early behavior problems ( $t = -1.99, p = .049$ ), and impulsivity ( $t = -4.769, p < .000$ ) (data not presented in the table; results for remaining items available on request).

**Table 6.** T-test for PCL-R, total and facet scores

Enforced treatment				
	No	Yes	t	p
n	53	55		
Facet 1	2.24 (2.0)	2.30 (1.8)	-.173	.863
Facet 2	4.77 (2.3)	5.76 (1.8)	-2.523	.013*
Facet 3	4.26 (2.5)	5.37 (2.4)	-2.361	.020*
Facet 4	2.82 (2.9)	3.72 (2.30)	-1.804	.074
Total	14.93 (8.0)	18.14 (6.5)	-2.293	.024*

Mean scores (SD) for PCL-R facets and total. \*  $p < .05$ .

## Discussion

Our results show that in Dutch forensic practice, where criminal responsibility is expressed on a five point scale, the presence of antisocial personality disorder did not lead to a ruling by the investigating team of complete absence of responsibility, but it did decrease responsibility up to a degree of diminished responsibility and lead to an advice for enforced forensic treatment. The role that a personality disorder plays cannot be completely attributed to comorbid Axis I psychopathology as personality disorders remained a predictor for diminished criminal responsibility in a regression analysis that corrected for other pathology. This means that Dutch forensic experts consider a personality disorder by itself to impair an individual's freedom of will. Although PCL-R total and facet scores did not differ between degrees of responsibility, results did show certain relevant items (Study 2). These three items concerned impulsivity and (ir)responsibility, suggesting that it is mostly deficiency in behavioral control that is taken into consideration by forensic experts in judging criminal accountability.

In contrast to the claims of Mei-Tal (2002) and Herpertz and Sass (2000), we did not find evidence that affective components such as lack of empathy or emotional deficiency are embraced by forensic experts as important factors in their assessments of criminal responsibility. Interestingly though, not only the presence of a personality disorder (Study 1) but also high PCL-R facet 2 and facet 3 scores (Study 2) turned out to be statistically significant predictors for advises of enforced treatment. Thus, despite the finding that affective deficiencies are not considered reason for diminished responsibility, such deficiencies are deemed sufficiently important by Dutch experts to indicate suspects' threat to society or dangerousness in combination with a need for treatment. In other words, both level of behavioral control and emotional functioning in individuals with a personality disorder seem to be relevant psychopathological conditions for forensic evaluations.

It should be noted that the results in both studies, although statistically significant and insightful into which factors experts take into account in their forensic pre-trial assessments, did not have very large effect-sizes. There was no one-to-one relationship between the presence of a personality disorder or characteristics of psychopathy on the one hand and diminished responsibility or advice for enforced treatment in a high security hospital on the other. It is therefore not warranted to conclude that the mere presence of a personality disorder or a high PCL-R score is sufficient cause for a ruling of diminished responsibility or an advice for enforced treatment. Nonetheless, the results of this study show that Dutch forensic clinicians take the diagnosis of a personality disorder or a high PCL-R score into careful consideration when making recommendations for diminished responsibility or the need for enforced treatment and for that reason should continue to be included in psychological and psychiatric assessments of suspects in criminal cases.

It should be noted that the Dutch legal system of deciding criminal responsibility differs substantially from that in other jurisdictions, and that the results in this study are difficult to generalize to the U.S., Canada, or some European countries. Also, while the inpatient records of the PBC, the specificity of this sample makes that the results might not generalize to outpatient forensic examinations in the Netherlands or other countries. Nonetheless, the PBC provided a sample with great variability in psychopathy scores and rates of personality disorders.





Chapter Six

*Prognostic factors  
for length of enforced  
forensic treatment  
with special focus on  
personality disorder*

## Abstract

The current study investigates the predictive value of personality disorder on treatment length, at first leave request and end of treatment, of offenders detained in high security psychiatric hospitals in the Netherlands. Data from digital databases of pre-trial assessment and treatment characteristics were merged for 536 male patients (mean age 34.0 years, SD = 10.5, range: 18.0-68.8) with a 76.5% personality disorder prevalence. Results of survival analyses, which also examined the predictive value of index offense, DSM-IV major mental disorder, substance abuse history, intellectual functioning, and psychiatric hospital as covariates, showed that personality disorder did not independently predict length of enforced treatment. A subgroup of patients with only personality disorder and no comorbid major mental disorder had a shorter treatment length. This study has isolated separate factors that can shorten or prolong treatment and can provide a focus for both policy and (clinical) decision makers in the enforced treatment process.<sup>5</sup>

## Enforced treatment

Under Dutch criminal law a person who commits a serious offense for which he cannot be held responsible, on account of mental retardation or pathological impairment of his mental abilities, is not punishable and can be sentenced to enforced treatment (known as *terbeschikkingstelling* or TBS) in special forensic psychiatric high security treatment hospitals, called Forensic Psychiatric Centers (FPCs). In cases in which the offender is considered partially unaccountable the treatment is preceded by a prison sentence, meant as punishment. The aim of enforced treatment is to protect society against a high risk of recidivism: in the short term by incarcerating the offender and in the long term by offender treatment (Van Gemmert, Van Schijndel, Gordeau, & Casanova, 2013). These hospitals aim towards the forensic psychiatric patient's gradual and safe return into society, known as "resocialization" (Koenraadt, Mooij, & Van Mulbregt, 2007).

The median enforced treatment length for forensic psychiatric patients had gradually been increasing over the years, from 7.3 years for patients who entered the system in 1990 (Dienst Justitiële Inrichtingen, 2011) to its peak at 10.6 years for patients who entered the system in 1998 (Dienst Justitiële Inrichtingen, 2015). A recent study shows that the median enforced treatment length for patients who entered the system in 2006 has decreased somewhat to 8.0 years (Dienst Justitiële Inrichtingen, 2015). Extensive treatment length is generally considered counterproductive to an effective resocialization process as well as overly costly (Verwaaijen & Polak, 2010). Hospitals face constant deliberation between preparing the patient for his or her return to society and the risk of reoffending society may face in doing so. Additional insight is needed into the conditions under which patients can responsibly return to society, because of the enormous public impact of any relapse into serious criminal behavior.

According to Brand, Van Emmerich and Raes (2009), research on the entire process of enforced treatment – which includes collecting data from and relaying it back to hospitals – can contribute to improved diagnoses and meaningful research, but also fewer incidents and less recidivism. Research into factors that predict length of treatment can increase the evidence base and the understanding of conditions under which patients can start their resocialization at an earlier stage. The present study intends to elaborate on current knowledge of personality disorder (PD), which is highly prevalent in Dutch hospitals (De Ruiter & Greeven, 2000; Hildebrand & De Ruiter, 2004), as a possible influence on

.....  
<sup>5</sup> Spaans, M., De Beurs, E., Rinne, T., & Spinhoven, P. (in press). Prognostic factors for length of enforced forensic treatment with special focus on personality disorder. *International Journal of Forensic Mental Health*.

treatment length of forensic psychiatric patients in hospitals in the Netherlands. It will do so by combining elaborate data from various judicial organizations that assess and treat mentally ill offenders.

### **Leave of forensic psychiatric patients**

A crucial part of the Dutch enforced forensic treatment is the leave process, during which the patient is granted authorized freedom to move outside the secure zone of the hospital. Authorized leave is meant to gradually prepare the patient for return to society, by examining if treatment advances made in the hospital hold up in the less protective outside environment and if the patient can handle an increase in freedom and responsibility (Nagtegaal, Van der Horst, & Schönberger, 2011; Bernstein, Nijman, Karos, Keulen-De Vos, De Vogel, & Lucker, 2012). It is not the forensic psychiatric patient but the hospital that can request leave on behalf of the patient, when the patient has shown sufficient progress during treatment. "Accompanied leave" is the first step and generally applied for after about two to three years of treatment (Waij, Spronk & Canton, 2010).

Since January 1st 2008 FPC's requests for leave are evaluated by the Advisory Board on Review of Leave from TBS (*Adviescollege Verloftoetsing TBS* or AVT) who, based on information from the hospital, such as standardized risk assessment instruments, consider the patient's possible risk to society during leave. Of the 1,868 decisions made by the Advisory Board in 2013, 1,640 (87.8%) leave requests were granted and only 136 (7.3%) were denied. For 92 (4.9%) requests, additional information was asked of the hospital before a decision could be made (AVT, 2014). According to Mevis (2011), an increasing fear of incidents with forensic psychiatric patients on leave has caused hospitals to prolong the first applications for leave of their patients, while at the same time the treatment directors of hospitals throughout the country are keen to speed up the leave process.

### **Influences on treatment length**

A number of factors affect the speed with which accompanied leave is requested for a forensic psychiatric patient. According to Van Kuijck (2011) hospitals differ greatly in the speed with which they apply for first leave, leading to differing numbers of patients per hospital who have never been on leave. This could be due to the fact that hospitals often specialize in types of offenders or disorders, such as patients who are considered violent, have major mental illnesses or below average intellectual functioning. Research has also shown that psychiatric problems that experts consider risk factors or contra-indications for accompanied leave are, in order of apparent priority, alcohol-related disorders, drug-related disorders, psychotic disorders, the Diagnostic Statistical Manual's (DSM; American Psychiatric Association, 2000) Cluster B personality pathology, and cognitive impairment (AVT, 2011). Nagtegaal and colleagues (2011) found that forensic psychiatric patients who had been convicted of sex offenses were increasingly less likely to leave hospitals.

When a forensic psychiatric patient has successfully passed through all leave stages, the hospital considers a patient for (conditional) end of enforced treatment. A patient whose enforced treatment has ended can still be supervised by Probation Services for a minimum of one year and a maximum of nine years. Research has shown that the length of supervision has gradually increased over the years from 343 days in 2009 to 537 days in 2013, which is most likely due to the increase in opportunities for aftercare and extended supervision during these years (Dienst Justitiële Inrichtingen, 2014).

## **Current study**

According to research by De Kogel and Nagtegaal (2006) most patients in forensic psychiatric hospitals in Canada, Sweden, England, and Wales suffer from severe psychiatric disorders such as schizophrenia and treatment is primarily pharmacological. In England, the treatment of personality disordered offenders has in recent years received more focus with the Dangerous and Severe Personality Disorder Program (DSPD) that was first piloted in 2001, and the ensuing Offender Personality Disorder Pathway (OPDP) program that was initiated in 2011. Both programs targeted treatment of offenders with a severe or complex PD, including high psychopathy scores, who pose a high risk of serious harm to others and of serious reoffending (Joseph & Benefield, 2012).

Dutch law takes into account the role PD can play in criminal behavior. Dutch forensic mental health experts consider the presence of a PD to diminish criminal responsibility and warrant enforced treatment (De Kogel & Nagtegaal, 2006; Spaans, Barendregt, Haan, Nijman, & De Beurs, 2011). However, longitudinal or prospective research on the relationship between PD and treatment length for forensic psychiatric patients in the Netherlands is limited.

A recent retrospective study on forensic psychiatric patients residing in one of the Netherlands' hospitals found that neither type of index offense nor psychiatric disorder influenced time to first accompanied leave (Ter Horst, Jessen, Bogaerts, & Spreen, 2015).

The current study investigated the predictive value of PD on treatment duration at first accompanied leave request and at end of enforced treatment of forensic psychiatric patients in forensic psychiatric high security treatment hospitals in the Netherlands by merging data from digital databases for the first time. To examine possible confounding, the effect of index offense, major mental disorder, substance abuse history, intellectual functioning and FPCs were also investigated. Taking into consideration the possible interaction effect on treatment length of comorbidity of multiple psychiatric disorders, relevant statistical analyses were repeated with a subgroup of patients that were diagnosed with only a PD and no other disorders.

## **Method**

### **Sample origin**

The Netherlands Institute of Forensic Psychiatry and Psychology (NIFP) is a national service of the Dutch Ministry of Justice. One of the NIFP's main tasks is to mediate between the prosecutor and independent pre-trial psychiatric and psychological experts who carry out assessments on criminal responsibility and the necessity of treatment. The NIFP's forensic mental health experts carry out around 6,250 assessments on adult suspects of crimes per year. Around 95% take place in a non-specialized forensic mental health setting (usually in a regular house of detention). The remaining 5% of forensic pre-trial evaluations are carried out in the Pieter Baan Center (PBC), the NIFP's forensic psychiatric observation clinic. These in-patient assessments take about seven weeks and cover the country's most severe crimes – that carry a possible prison sentence of four years or more – and/or the most complex suspected psychopathology.

In order to obtain the study sample for which additional data would be requested from partner organizations, court records of all 16,409 suspects of crimes assessed by the NIFP between January 1, 2003 and December 31, 2006 were investigated. Because of the low number of female defendants assessed by the NIFP during this time (7.9%), it was decided to include only male subjects in the study. Of 680 verdicts of enforced treatment for an indefinite period of time for male adult suspects who were assessed by the NIFP before their enforced treatment began, only those cases were selected with complete digital NIFP records for presence or absence of a PD diagnosis, and whose enforced treatment had started as of September 1, 2014. This resulted in a final study sample of 536 convicted male offenders who had been ordered to undergo enforced treatment in a high security hospital.

### **Databases**

In addition to the data available in the NIFP's Forensic Registration and Information System (FRIS) and PBC databases, data for these 536 cases was requested from a partner organization in the judicial chain: the Monitor Information TBS (MITS) database from the Custodial Institutions Agency (*Dienst Justitiële Inrichtingen*; DJI).

All diagnostic data from the FRIS database are classified according to the DSM-IV(-TR) (American Psychiatric Association, 1994; American Psychiatric Association, 2000), which were the most current versions of the DSM at the time the data was collected and the reason for the multi-axial approach to data analysis. PDs are listed as Clusters A, B, and C, and PD Not Otherwise Specified (NOS). Independent forensic mental health experts use various diagnostic instruments to establish the presence of major mental disorders, substance abuse history, PDs, and intellectual functioning. All classifications in the present study, including those for PDs, were clinician-based, supported by self-report questionnaire scores. Examples of validated self-report instruments that were used as possible indications of PDs are the Dutch version of the Minnesota Multiphasic Personality Inventory-2 (MMPI-2; Hathaway & McKinley, 1989), the Edwards Personal Preference Schedule (EPPS; Edwards, 1959) and the Dutch Personality Questionnaire (NPV; Barelds, Luteijn, Van Dijk, & Starren, 2007). The inpatient assessment setting uses semi-structured diagnostic interviews where possible, such as the Structured Interview for DSM-IV Personality (SIDP-IV; Pfohl, Blum, & Zimmerman, 2006).

### **Safeguards of anonymity**

All data was received anonymously, meaning that no names were ever exchanged between the organizations. Data was matched based on the registration numbers that correspond to court cases. Also, to guarantee anonymity of the hospitals represented in this study, each hospital's name was replaced with a number only known to the authors of this study.

### **Variables**

General characteristics were date and country of birth, date of admission to the hospital, name of the hospital, type of index offence (property and fraud, property and violence, (attempted) violence, (attempted) sex offense, (attempted) murder or manslaughter, and fire setting), diagnostic information from pre-trial NIFP assessment such as DSM-IV major mental disorders (such as disorders usually first diagnosed in infancy, childhood, or adolescence, affective disorder, paraphilia, and psychotic symptoms), PDs (Clusters A, B, and C and PD NOS), level of intellectual functioning (below average, average, and above average), and substance abuse history divided into alcohol, cannabis, or other/poly-substance abuse, conclusion on criminal responsibility, and advice on enforced treatment. Information from the

hospitals included the start date of enforced treatment, the dates on which the accompanied leave was first requested and the date on which enforced treatment was officially terminated, if applicable.

### **Statistical analysis**

Treatment length was operationalized at two separate events: at first accompanied leave request and at end of enforced treatment. This was done by calculating the time (in years) between the start date of enforced treatment and the date on which accompanied leave or termination of treatment was first requested by the hospital. Cohen's *d* coefficient was calculated for the differences in treatment length at both first accompanied leave request and at end of enforced treatment between each category of the independent variable and the reference category.

Cox regression survival analyses were carried out to explore the predictive effect of PD, index offense, major mental disorder, substance abuse history, and intellectual functioning on time to first accompanied leave request and at the end of enforced treatment. First, the predictive value of each of the individual predictor variables was analyzed separately (Model 1). Next, separate baseline hazard functions were fitted within different strata (i.e., hospitals with high versus low average time to first accompanied leave request and to end of treatment, respectively) while still modeling the effects of the predictors as a single set of common effects across strata (Model 2). In the case of more than one significant result in Model 2, a stratified Cox regression model (Model 3) including all significant predictors at  $p < .10$  according to Model 2 was executed. Linear regression analyses were carried out on a small subset of patients who received a diagnosis of PD alone.

All analyses were carried out with Statistical Package for the Social Sciences (SPSS) versions 17 and 22. In Cox regression analyses, SPSS creates its own dummy categories. The first category within each independent variable was assigned as the reference category using the Indicator contrast, except for intellectual functioning for which average intellectual functioning was selected as the reference category.

## **Results**

### **Sample characteristics**

The data sample consisted of 536 men convicted of one or more crimes that carry a possible prison sentence of four years or more. Of these assessments, 428 (79.9%) were carried out during the participant's stay in a regular house of detention and the remaining 108 (20.1%) were carried out during a seven-week admission to the Pieter Baan Center. The mean age at the time of all 536 assessments was 34.0 years ( $SD = 10.5$ , range: 18.0–68.8 years).

The date of the official start of the enforced treatment – the date of early release from prison for those who were deemed diminished responsible and the date on which the court sentence was finalized for those found fully not responsible – ranged from May 2003 to February 2013. The average length of time between the NIFP assessment and the start of enforced treatment was 1.4 years ( $SD = 1.2$  years, range: 0.1–7.5 years). The majority of the treatments started in 2006 ( $n = 132$ ; 24.6%), 2005 ( $n = 118$ ; 22.08%), and 2007 ( $n = 115$ ; 21.5%). It is important to note that new regulations in 2010 prohibited prisoners from starting enforced treatment after having carried out only a third of their prison sentence and increased the minimum of time spent in prison to two thirds of the sentence.

PDs were present in 410 defendants (76.5% of the total sample), according to the NIFP. Table 1 shows

the prevalence of the categories of PD in the study sample and of the independent variables index offense, major mental disorder, substance abuse history, intellectual functioning, and the hospital that first requested accompanied leave or from which the enforced treatment was terminated. Comorbidity of major mental disorder and PD was present in 249 defendants (46.5% of the total sample).

Table 1 also shows average treatment length for each category at time of first accompanied leave request ( $n = 446$ ) and at end of treatment ( $n = 81$ ), as well as the value for Cohen's  $d$  for the difference with the reference category within each independent variable. Given the differences in treatment length between FPCs (range for first accompanied leave request: 2.77–3.81 years; range for end of treatment: 3–8 years) carrying out further statistical analyses that were stratified for hospitals with high versus low average time to first accompanied leave request and to end of treatment, respectively, was considered warranted. The index offenses that led to NIFP assessments are categorized according to the seriousness index of Van Kordelaar (2003). The higher the category number, the more serious the offense. Van Kordelaar makes no distinction between an attempted offense and an executed offense – it is the general nature of the crime that indicates seriousness. Not all categories of the seriousness index were represented in the current study.

**Table 1.** Prevalence of independent variables and average treatment length

	Treatment length (in years) at							
	Prevalence $n = 536$		Accompanied leave $n = 446$			End of treatment $n = 81$		
<i>Personality disorder</i>	<i>n</i>	<i>%</i>	<i>n</i>	<i>M (SD)</i>	<i>Cohen's d</i>	<i>n</i>	<i>M (SD)</i>	<i>Cohen's d</i>
No personality disorder	126	23.5	102	3.58 (1.39)	(ref. cat.)	19	6.15 (1.84)	(ref. cat.)
Personality disorder Cluster A	13	2.4	12	3.34 (1.43)	-.17	5	6.99 (1.42)	.51
Personality disorder Cluster B	224	41.8	190	3.50 (1.68)	-.05	30	6.19 (1.74)	.02
Personality disorder Cluster C	19	3.5	17	3.27 (1.15)	-.24	1	7.05 (-)	-
Personality disorder NOS	154	28.7	125	3.34 (1.64)	-.16	26	6.54 (1.72)	.22
<i>Index offense</i>	<i>n</i>	<i>%</i>	<i>n</i>	<i>M (SD)</i>	<i>Cohen's d</i>	<i>n</i>	<i>M (SD)</i>	<i>Cohen's d</i>
Property and fraud	36	6.7	28	3.88 (2.04)	(ref. cat.)	6	6.10 (1.69)	(ref. cat.)
(Attempted) violence	93	17.4	76	3.38 (1.53)	-.28	18	5.97 (1.94)	-.07
Property and violence	31	5.8	25	3.32 (1.53)	-.31	8	6.35 (1.71)	.15
(Attempted) sex offense	122	22.8	102	3.90 (1.71)	.01	6	7.46 (1.52)	.85

Treatment length (in years) at								
	Prevalence <i>n</i> = 536		Accompanied leave <i>n</i> = 446			End of treatment <i>n</i> = 81		
(Attempted) murder/manslaughter	201	37.5	172	3.22 (1.33)	-.38	32	6.60 (1.30)	.33
Fire setting	49	9.1	39	3.35 (1.75)	-.28	10	5.55 (2.40)	-.26
Other / unknown	4	0.7	4	2.70 (1.54)	-	1	-	-
<i>Axis I diagnoses</i>	<i>n</i>	%	<i>n</i>	<i>M</i> ( <i>SD</i> )	<i>Cohen's d</i>	<i>n</i>	<i>M</i> ( <i>SD</i> )	<i>Cohen's d</i>
No Axis I disorder	154	28.7	126	3.19 (1.46)	(ref. cat.)	25	6.16 (1.64)	(ref. cat.)
Disorders usually first diagnosed in infancy, childhood, or adolescence	67	12.5	56	3.66 (1.78)	.29	11	6.14 (1.68)	-.01
Affective disorder	38	7.1	32	3.29 (1.19)	.08	8	7.27 (.94)	.83
Paraphilia	60	11.2	51	3.58 (1.53)	.26	4	7.32 (1.58)	.72
Psychotic symptoms	143	26.7	118	3.71 (1.65)	.33	21	6.72 (1.92)	.31
Other	44	8.2	36	3.35 (1.65)	.10	6	5.80 (1.76)	-.21
Unknown (not assessed)	30	5.6	27	3.32 (1.62)	-	6	-	-
<i>Substance use</i>	<i>n</i>	%	<i>n</i>	<i>M</i> ( <i>SD</i> )	<i>Cohen's d</i>	<i>n</i>	<i>M</i> ( <i>SD</i> )	<i>Cohen's d</i>
No (diagnosis) substance use	178	33.2	141	3.46 (1.54)	(ref. cat.)	29	6.34 (1.65)	(ref. cat.)
Alcohol	96	17.9	79	3.21 (1.56)	-.16	14	5.41 (1.53)	-.58
Cannabis	65	12.1	52	3.59 (1.59)	.08	11	6.51 (1.38)	.11
Other/Poly	150	28.0	132	3.53 (1.64)	.04	18	6.35 (1.95)	.01
Unknown (not assessed)	47	8.8	42	3.52 (1.54)	-	9	-	-
<i>Intellectual functioning(IF)</i>	<i>n</i>	%	<i>n</i>	<i>M</i> ( <i>SD</i> )	<i>Cohen's d</i>	<i>n</i>	<i>M</i> ( <i>SD</i> )	<i>Cohen's d</i>
Below average IF (<90)	222	41.4	182	3.58 (1.60)	-.12	27		-.02
Average IF (90-110)	205	38.2	168	3.39 (1.50)	(ref. cat.)	30	6.46 (1.36)	(ref. cat.)

	Treatment length (in years) at							
	Prevalence <i>n</i> = 536		Accompanied leave <i>n</i> = 446			End of treatment <i>n</i> = 81		
Above average IF (>110)	42	7.8	40	3.05 (1.48)	.23	9	5.98 (2.04)	.28
Unknown (not assessed)	67	12.5	56	3.56 (1.77)	-	15	-	-
FPC	<i>n</i>	%	<i>n</i>	<i>M</i> ( <i>SD</i> )	<i>Cohen's d</i>	<i>n</i>	<i>M</i> ( <i>SD</i> )	<i>Cohen's d</i>
Administrative department	-	-	0	(-)	-	1	3.00 (-)	-
FPC 1	-	-	42	3.62 (1.60)	-	8	6.12 (1.94)	-
FPC 2	-	-	49	3.81 (1.62)	-	3	5.73 (1.54)	-
FPC 3	-	-	47	2.94 (1.39)	-	10	6.44 (1.69)	-
FPC 4	-	-	70	3.77 (1.56)	-	15	6.68 (1.10)	-
FPC	<i>n</i>	%	<i>n</i>	<i>M</i> ( <i>SD</i> )	<i>Cohen's d</i>	<i>n</i>	<i>M</i> ( <i>SD</i> )	<i>Cohen's d</i>
FPC 5	-	-	36	2.98 (1.57)	-	6	6.91 (1.21)	-
FPC 6	-	-	27	3.77 (1.89)	-	7	6.39 (1.77)	-
FPC 7	-	-	1	3.60 (-)	-	2	3.00 (1.42)	-
FPC 8	-	-	21	3.57 (1.01)	-	-	-	-
FPC 9	-	-	11	3.77 (1.43)	-	5	7.54 (1.62)	-
FPC 11	-	-	70	3.69 (1.55)	-	11	7.12 (1.27)	-
FPC 12	-	-	14	3.81 (2.25)	-	1	5.49 (-)	-
FPC 13	-	-	53	2.77 (1.34)	-	11	5.36 (2.02)	-
FPC 14	-	-	5	2.98 (1.02)	-	1	8.00 (-)	-

Note. NOS = Not Otherwise Specified. Ref.cat. = reference category.

## Accompanied leave

As of September 1, 2014, hospitals had applied for at least one accompanied leave for 446 of the 536 patients (83,2%). Length of treatment at time of first application for accompanied leave was calculated with the start date of treatment and the date of first application and varied from 0.73-9.97 years with a mean of 3.46 years (SD = 1.58 years). In order to assess the independent effect of the various categories of PD on treatment length at first accompanied leave application for all 536 participants, Cox regression analyses were carried out in three phases.

First, a univariate Cox regression analysis was carried out with PD as predictor and treatment length at first accompanied leave application as outcome for all 536 participants. To determine whether index offense, major mental disorder, substance abuse history, and intellectual functioning might also influence treatment length at first accompanied leave application, univariate Cox regression analysis was carried out on these variables separately as well (Model 1). Hazard ratios and 95% confidence intervals for all independent variables are shown in Table 2.

Results of the Cox regression analyses were not significant for PD. Due to the high prevalence of PD NOS in the study sample (n = 154; 28.7%), the Cox regression analyses were repeated without this category, with the absence of PD still as the reference category. These analyses also showed no significant results. Results of the Cox regression analyses were also not significant for major mental disorder and substance abuse history as predictors. Results for index offense showed significant results for an index offense of (attempted) murder/manslaughter, HR = 1.57, 95% CI [1.05, 2.34], p = .028, suggesting that the time until first accompanied leave application was shorter for this index offense than for the reference category of property and fraud. Results of the Cox regression analysis for intellectual functioning showed significant results for above average intellectual functioning, HR = 1.60, 95% CI [1.14, 2.27], p = .007, suggesting that the time until first accompanied leave application was shorter for above average intellectual functioning than for the reference category of average intellectual functioning.

Subsequent Cox regression analyses for time to first accompanied leave request of the individual predictor variables with hospital as stratum (Model 2) also showed no significant results for PD, major mental disorder, and substance abuse history. The same was true for analysis without the PD NOS category. There was no longer a significant effect for above average intellectual functioning.

The significant effect on time to first accompanied leave application remained for an index offense of (attempted) murder/manslaughter, with a shorter time to first accompanied leave application compared to the reference category of property and fraud, HR = 1.91, 95% CI [1.27, 2.88], p = .002. There was also a significant effect for an index offense of (attempted) violence on time to first accompanied leave application, with a shorter time to first accompanied leave application compared to the reference category of property and fraud, HR = 1.71, 95% CI [1.10, 2.67], p = .018.

**Table 2.** Results of Cox regression analyses of time to first accompanied leave request in three phases (n=536)

	Model 1		Model 2	
	Univariate		Univariate with stratified FPC	
<i>Personality disorder</i>	HR	95% CI	HR	95% CI
No personality disorder (ref.cat.)	-	-	-	-
Personality disorder Cluster A	1.35	[.74, 2.45]	1.12	[.61, 2.04]
Personality disorder Cluster B	1.16	[.91, 1.47]	.96	[.75, 1.22]
Personality disorder Cluster C	1.47	[.88, 2.46]	1.17	[.70, 1.97]
Personality disorder NOS	1.12	[.86, 1.45]	.92	[.70, 1.20]
<i>Index offense</i>	HR	95% CI	HR	95% CI
Property and fraud (ref.cat.)	-	-	-	-
(Attempted) violence	1.26	[.82, 1.95]	1.71**	[1.10, 2.67]
Property and violence	1.29	[.75, 2.21]	1.54	[.89, 2.68]
(Attempted) sex offense	1.16	[.76, 1.76]	1.29	[.84, 1.97]
(Attempted) murder/manslaughter	1.57**	[1.05, 2.34]	1.91**	[1.27, 2.88]
Fire setting	1.19	[.73, 1.94]	1.41	[.86, 2.31]
<i>Major mental disorders</i>	HR	95% CI	HR	95% CI
No major mental disorder (ref.cat.)	-	-	-	-
Disorders usually first diagnosed in infancy, childhood, or adolescence	.90	[.66, 1.23]	.79	[.57, 1.09]
Affective disorder	1.01	[.69, 1.49]	1.01	[.69, 1.49]
Paraphilia	.95	[.69, 1.32]	.88	[.63, 1.22]
Psychotic symptoms	.81	[.63, 1.05]	1.03	[.80, 1.34]
Other	.96	[.66, 1.38]	.98	[.68, 1.42]
<i>Substance use</i>	HR	95% CI	HR	95% CI
No substance use (ref.cat.)	-	-	-	-
Alcohol	1.14	[.87, 1.50]	1.07	[.81, 1.41]
Cannabis	.93	[.67, 1.27]	1.06	[.83, 1.34]
Other/Poly	1.21	[.95, 1.54]	.94	[.68, 1.29]
<i>Intellectual functioning (IF)</i>	HR	95% CI	HR	95% CI
Below average IF (<90)	.96	[.78, 1.19]	.87	[.70, 1.07]
Average IF (90-110; ref.cat.)	-	-	-	-
Above average IF (>110)	1.60**	[1.14, 2.27]	1.18	[.83, 1.68]

Note. NOS = Not Otherwise Specified. HR = Hazard Ratio. CI = Confidence Interval. Ref.cat. = reference category. IF = Intellectual functioning. \*\* = Significant at  $p < .05$ . \* =  $.05 > p < .10$ .

### **Explorative analysis: Pure PD**

It was considered a possibility that the enforced treatment could be aimed at multiple psychiatric disorders at the same time, such as PD, a major mental disorder, *and* substance abuse. For that reason, those forensic psychiatric patients were selected that had a diagnosis of PD only, which was the case for 62 of the 410 patients with a PD. Of these 62 patients (15.1%), one patient (1.6%) had a Cluster A PD, 31 patients (50%) had a Cluster B PD, four patients (6.5%) had a Cluster C PD, and 26 patients (41.9%) received the classification of PD NOS. A Cox regression analysis with all categories of PD indicated a shorter time to first accompanied leave application for this group than for the group with comorbidity, HR = 1.39, 95% CI [1.05, 1.84],  $p = .021$ . Excluding the PD NOS category from this analysis led to a non significant result.

The possibility was considered that, when applying for leave for the first time, decision-makers take into account the prison sentence carried out before enforced treatment by patients found partially criminally responsible. To first determine whether patients with PD only were more likely to be found partially responsible, and therefore had already served a period of time in custody, their degree of criminal responsibility according to the NIFP was divided into three categories out of the five that were used at the time of data collection: not responsible, partially responsible and fully responsible. Inspection of frequencies showed that 56 of these patients (90.3%) were found partially responsible. Next, a linear regression analysis was carried out to determine whether the time between the NIFP assessment and the start of enforced treatment – as a proxy for time spent in prison before treatment – influenced the time until leave was first requested for the PD only patients. Results were not significant, indicating no significant relationship between the two variables,  $R^2 = .01$ ,  $F(1, 60) = .62$ ,  $p = .43$ .

### **End of enforced treatment**

In the data provided by DJI in January of 2014, the enforced treatment had been ended by the judge (for reasons other than deportation/extradition, legal error, a new verdict, or death of the patient) for 81 individuals (15.1% of the total sample). The advice of the hospital to end the enforced treatment was followed in 44 cases (54.3%) and the advice of the hospital to continue treatment was opposed in 25 cases (30.9%). Twelve patients (14.8%) were transferred to a psychiatric hospital. Total length of treatment was calculated with the official start and end dates of treatment and varied from 2.00-10.12 years with a mean of 6.35 years (SD = 1.72 years). The earliest release date was February 2007.

In order to assess the independent effect of the various categories of PD, index offense, major mental disorder, substance abuse history, and intellectual functioning on total treatment length at time of release for all 536 participants, Cox regression analyses according to Models 1 and 2 were repeated for total length of treatment. Hazard ratios and 95% confidence intervals for all independent variables are shown in Table 3.

In Model 1 a predictors at  $p < .10$  for shorter total treatment length was found for Cluster A PD, HR = 2.43, 95% CI [.90, 6.51],  $p = .078$ . The same analysis without the category of PD NOS yielded an effect at  $p < .10$  for Cluster A PD, HR = 2.50, 95% CI [.93, 6.70],  $p = .069$ . Results for index offense showed significantly longer treatment length for an index offense of (attempted) sex offense, HR = .26, 95% CI [.08, .80],  $p = .018$ , and for disorder paraphilia, HR = .33, 95% CI [.11, .94],  $p = .038$ . Results for substance abuse and intellectual functioning were not significant.

According to results of the Cox regression analyses in Model 2 with hospital as strata and all predictors at  $p < .10$  according to Model 1, substance abuse history and intellectual functioning remained not significant while the effect at  $p < .10$  of PD was no longer observed in both analyses with and without PD NOS. The significantly longer total treatment length remained for an index offense of (attempted)

sex offense, HR = .21, 95% CI [.06, .71],  $p = .012$ . The significant effect of paraphilia on total treatment length found in Model 1 was just below  $p < .10$  in Model 2, HR = .40, 95% CI [.14, 1.16],  $p = .092$ . Finally, a stratified Cox regression model including all predictors at  $p < .10$  according to Model 2 (i.e., index offense and major mental disorder) was executed. This third model showed a significant effect across strata for the index offense of (attempted) sex offense on total treatment length, HR = .11, 95% CI [.02, .58],  $p = .010$ . The effect of paraphilia was no longer significant after index offense was also included into the model.

### Explorative analysis: Pure PD

Again, a Cox regression analysis was carried out within the group of 410 patients with PD of whom 62 patients only had a PD diagnosis and no other psychiatric comorbidity (see above for classifications). Results with time to enforced treatment ending as outcome showed a significantly shorter total treatment length for patients with only PD than for those with comorbidity of disorders, HR = 1.99, 95% CI [1.14, 3.49],  $p = .016$ . Excluding the PD NOS category from this analysis again led to a non significant result.

Again, taking into account the possible influence on decision-makers of the time elapsed between the NIFP assessment and the start of enforced treatment – as a proxy for time spent in prison before treatment – on total treatment length for the PD only patients, a linear regression analysis was carried out with total treatment length as dependent variable. Results were significant,  $R^2 = .38$ ,  $F(1, 60) = 37.4$ ,  $p < .00$ ;  $\beta = -.62$ ,  $p < .001$ , indicating a negative relationship between the time spent in prison before treatment and total treatment length.

**Table 3.** Results of Cox regression analyses for end of enforced treatment in three phases

	Model 1 Univariate		Model 2 Univariate with stratified FPC		Model 3 Index offense, major mental disorder and stratified FPC	
	HR	95% CI	HR	95% CI	HR	95% CI
<i>Personality disorder</i>						
No personality disorder (ref.cat.)	-	-	-	-	-	-
Personality disorder Cluster A	2.43*	[.90, 6.51]	2.16	[.80, 5.85]	-	-
Personality disorder Cluster B	.85	[.48, 1.50]	.78	[.44, 1.39]	-	-
Personality disorder Cluster C	.40	[.05, 3.01]	.78	[.10, 5.88]	-	-
Personality disorder NOS	1.14	[.63, 2.05]	.78	[.43, 1.44]	-	-
<i>Index offense</i>						
Property and fraud (ref.cat.)	-	-	-	-	-	-
(Attempted) violence	1.08	[.43, 2.73]	.75	[.30, 1.90]	.93	[.35, 2.49]
Property and violence	1.76	[.61, 5.09]	.85	[.29, 2.51]	.80	[.26, 2.49]
(Attempted) sex offense	.26**	[.08, .80]	.21**	[.06, .71]	.11**	[.02, .58]
(Attempted) murder/ manslaughter	.98	[.41, 2.35]	.89	[.37, 2.14]	.92	[.37, 2.28]
Fire setting	1.00	[.36, 2.75]	1.48	[.53, 4.13]	1.35	[.45, 4.09]

	Model 1 Univariate		Model 2 Univariate with stratified FPC		Model 3 Index offense, major mental disorder and stratified FPC	
<i>Major mental disorder</i>	HR	95% CI	HR	95% CI	HR	95% CI
No major mental disorder (ref.cat.)	-	-	-	-	-	-
Disorders usually first diagnosed in infancy, childhood, or adolescence	.91	[.45, 1.86]	.82	[.39, 1.73]	1.08	[.50, 2.31]
Affective disorder	1.15	[.52, 2.55]	.74	[.33, 1.66]	.69	[.30, 1.58]
Paraphilia	.33**	[.11, .94]	.40*	[.14, 1.16]	2.73	[.50, 14.97]
Psychotic symptoms	.67	[.37, 1.20]	.73	[.40, 1.31]	.68	[.36, 1.28]
Other	.74	[.30, 1.80]	1.28	[.51, 3.18]	1.45	[.57, 3.70]
<i>Substance use</i>	HR	95% CI	HR	95% CI	HR	95% CI
No substance use (ref.cat.)	-	-	-	-	-	-
Alcohol	.86	[.46, 1.63]	1.57	[.80, 3.07]	-	-
Cannabis	.89	[.45, 1.79]	.86	[.42, 1.73]	-	-
Other/Poly	.65	[.36, 1.18]	.74	[.41, 1.35]	-	-
<i>Intellectual functioning (IF)</i>	HR	95% CI	HR	95% CI	HR	95% CI
Below average IF (<90)	.75	[.45, 1.27]	1.23	[.71, 2.14]	-	-
Average IF (90-110; ref.cat.)	-	-	-	-	-	-
Above average IF (>110)	1.68	[.80, 3.55]	.77	[.35, 1.71]	-	-

Note. NOS = Not Otherwise Specified. HR = Hazard Ratio. CI = Confidence Interval. Ref.cat. = reference category. \*\* = Significant at  $p < .05$ . \* =  $.05 > p < .10$ .

## Discussion

The current study prospectively investigate the predictive value of personality disorder, present in 76.5% of a study sample of 536 forensic psychiatric patients from various forensic psychiatric high security treatment hospitals, on treatment length of enforced treatment at first accompanied leave request and at end of enforced treatment. To examine possible confounding, the effect of index offense, major mental disorder, substance abuse history, intellectual functioning, and hospitals were also investigated.

Statistical analysis of merged data from the NIFP and DJI databases showed no independent predictive value of PD for both treatment length at first accompanied leave request and total treatment length at the end of treatment. Multivariate analyses showed that an index offense of (attempted) violence and (attempted) murder/manslaughter both independently predicted shorter treatment length at first accompanied leave request across different strata for hospitals. When examining total treatment length at end of treatment, an index offense of (attempted) sex offense independently predicted a longer enforced treatment length, again across different strata for hospitals. These findings are in line with those of Nagtegaal and colleagues (2011) who found that forensic psychiatric patients who had been convicted of sex offenses were increasingly less likely to leave the hospitals. (Attempted) violence

is among the least severe offenses according to Van Kordelaar (2003). These offenders are apparently consequently considered by forensic mental health experts to present a lower risk of incidents during accompanied leave than other offenders in this study.

Post-hoc explorative analyses on a subgroup of patients with a diagnosis of *only* PD showed that this group had a shorter treatment length at both first accompanied leave request and end of enforced treatment than personality disordered patients who also had comorbid major mental disorder. Analysis of the influence of the time elapsed between the NIFP assessment and the start of enforced treatment – as a proxy for time spent in prison before treatment – on total treatment length for the PD only patients showed a negative relationship between the time spent in prison before treatment and total treatment length. It must be noted that both explorative analyses were carried out on a small subset of patients.

### **Strengths and limitations**

The present study combined data from different sources available in the Dutch judicial system and over a range of hospitals. Combining diagnostic information that is available at the time of imposing enforced treatment to predict the actual length of treatment later on yields important data regarding what can be expected of treatment for different groups of forensic psychiatric patients. On the other hand, in some cases more than ten (treatment) years had passed since the NIFP's diagnosis. Due to lack of data on a national level, the current study was unable to incorporate patient characteristics from the treatment itself.

The use of digital databases introduced some limitations to the current study. Besides the possibility of human error occurring when data was manually entered into both databases, there is usually little room for nuance in a digital database. There were substantial differences in number of patients and treatment length between hospitals. To control for these effects, stratified analyses were performed in order to model the effects of the predictors as a single set of common effects across strata.

The NIFP database was not able to distinguish between the various DSM-IV(-TR) PDs, only between Cluster A, B, and C and NOS. Also, the NIFP database was only able to give one PD diagnosis per person, so the current study was not able to investigate possible comorbidity within PDs or differences in severity of (personality) disorders. Another limitation associated with the use of the NIFP's digital database is that there is no information regarding the assessment instruments used to reach or support expert diagnoses on PD, major mental disorder, substance abuse history, and intellectual functioning. With regard to the diagnosis of PD, outcomes have been shown to vary according to assessment methods, such as self-report vs. interview methods (Blackburn, Donnelly, Logan, & Renwick, 2004; Guy, Poythress, Douglas, Skeem & Edens, 2008; Whyte, Fox, & Coxell, 2006), and legal contexts such as pre- and post-conviction (Cima et al., 2003).

The prevalence of PD NOS in this study is high, at 28.7% of the study sample. This is a common finding (Baillie & Lampe, 1998; Berhuis, 2014; Pagan, Oltmanns, Whitmore, Turkheimer, 2005; Verheul & Widiger, 2004; Zimmerman, Rothschild, & Chelminski, 2005). It could be that forensic mental health experts did not find the diagnostic categories of the DSM-IV(-TR) adequate to cover the personality pathology they considered in their clinical assessment, or that the experts felt the patient met the general diagnostic criteria for a PD but not the criteria for one of the individual disorders (Verheul & Widiger, 2004). It could also be the case that forensic mental health experts used PD NOS when two or more specific diagnoses could have been made (Verheul & Widiger, 2004), or that they prefer a more dimensional approach towards the classification of PDs, rather than a categorical one (Zimmerman, Rothschild, & Chelminski, 2005). On the other hand, several authors have doubted the validity of the DSM-IV(-TR) to diagnose PDs

(Bernstein, Iscan, & Maser, 2007; Clark, Watson, & Reynolds, 1995; Morey et al., 2012; Mulder, Newton-Howes, & Tyrer, 2011; Widiger, Livesley, & Clark, 2009). The high prevalence of PD NOS could cloud results as a patient's symptoms can transcend several PD categories or even the three clusters. Therefore, the high prevalence of PD NOS was taken into account in the current study by repeating relevant statistical analyses without this category of patients.

There was no digital data available on diagnoses at the time of first accompanied leave request and at end of treatment, or any measurement of within-treatment change. It is not possible to draw conclusions on treatability, and to tell whether the disorder diagnosed by the NIFP and its concurrent risk of recidivism was successfully treated or whether other forensically relevant characteristics or risk factors were managed during treatment. Also, as post-treatment outcome data, such as on possible reoffending after release, was not incorporated in the current study, it is not clear whether the entire resocialization process associated with enforced treatment was successful, i.e. if all released patients have been abiding by the law.

### **Suggestions for future research**

The current study has shown that the factors involved in the prediction of a successful resocialization process of disordered offenders are complex and intertwined. It can be considered a first attempt at distilling these factors from rough, unrefined digital data. Merging of existing data from various sources revealed inconsistencies in the level of detail of the digital data per person, resulting in limits to not only the comparability of the separate databases but also the generalizability of the current findings. In order to carry out further research, an increase in (detail of) digital registration of patient, treatment, and risk characteristics along the course of treatment and at its end is imperative. As mentioned earlier, Brand and colleagues (2009) advocate research on the entire process of enforced treatment by collecting data from and relaying it back to hospitals. Measurement feedback systems like Routine Outcome Monitoring (ROM) – measuring the patient's condition at fixed times and time periods aimed at evaluation and possible adjustment of treatment – are a way to gain insight into the effectiveness of the enforced treatment and risk reduction. Carlier's literature study (2010) on the effectiveness of ROM found a significant positive impact on diagnosis, monitoring of treatment and the physical and mental health status of patients. Because the majority of ROM is based on self-report instruments, which have been found to be less reliable in a forensic mental health setting (Spaans, Barendregt, Muller, De Beurs, Nijman, & Rinne, 2009; Spaans, Rinne, De Beurs, & Spinhoven, 2015), alternatives such as self-report instruments that include response bias indicators or clinician-administered instruments are recommended.

## **Conclusion**

To conclude, the present study combined national data on disordered offenders of serious crimes to longitudinally investigate the effect of personality disorder on first accompanied leave request and total length of enforced treatment across all treatment hospitals in the Netherlands. Characteristics that independently lead to shorter enforced treatment length at first accompanied leave request were an index offense of (attempted) murder/manslaughter and an index offense of (attempted) violence. Sex offenses lead to longer total treatment length. Although PD did not independently predict treatment length for the total sample, shorter treatment length was found in a subgroup of patients with only PD.



Chapter Seven  
*Discussion*

The present dissertation investigated personality pathology in a forensic mental health setting, largely using a database of results on assessment instruments acquired after a thorough revision of the assessment process in the Pieter Baan Center in the Netherlands. Chapter Two presented a literature review and meta-analytic study of the prevalence of the self-reported levels of the maladaptive personality traits anger, aggression, hostility, antisocial traits, psychopathy, and impulsivity in detained populations, compared to normal populations. In Chapter Three a study was presented into the utility of the Dimensional Assessment of Personality Pathology-Short Form (DAPP-SF) as a screener for personality disorder in a forensic psychiatric hospital that carries out pre-trial evaluations on suspects of serious crimes. The aim of the study was to ascertain whether the self-report instrument could be used as the first step in a two-step approach for an efficient assessment of personality psychopathology, and whether it could correctly determine who should and should not undergo a standardized (semi-) structured diagnostic interview to verify the presence of the disorder. Chapter Four presented a study on the use of another self-report instrument in a forensic sample: the Minnesota Multiphasic Personality Inventory-2 (MMPI-2). It assessed whether a cluster analysis of MMPI-2 profiles produces a multitude of distinct personality profiles when assessing personality traits in a known heterogeneous population of pretrial criminal defendants of serious crimes in a forensic psychiatric observation hospital. The study presented in Chapter Five investigated how Dutch experts consider personality disorder, compared to other psychiatric conditions, in their recommendations regarding criminal responsibility and in their advice on the necessity of enforced treatment in a high security hospital. It also investigated how they consider various aspects of psychopathy as measured by the Psychopathy Checklist-Revised (PCL-R) into these judgments. Lastly, Chapter Six examined the enforced treatment of personality disordered criminal offenders in high security hospitals and the prospective relationship between personality disorder and treatment duration at first leave request and end of treatment. This study also investigated possible confounding effects of index offense, Axis 1 disorder, substance abuse history and intellectual functioning.

## Chapter Two: Prevalence of personality pathology in forensic mental health settings

Before this dissertation study commenced, it was already clear that personality pathology was associated with criminal offending and violence towards others and highly prevalent in forensic settings (Fazel & Danesh, 2002), especially DSM Cluster B personality disorders (De Ruiter & Greeven, 2000; Hildebrand & De Ruiter, 2004; Timmerman & Emmelkamp, 2001) and related maladaptive personality traits such as aggression or hostility, poor impulse control, sensation seeking, and lack of empathy as assessed with self-report assessment instruments (Boccaccini et al., 2010; Cunradi et al., 2009; Dolan & Blackburn, 2006; Norlander & Eckhardt, 2005; Shechory et al., 2011; Walters, 2007). Doubts about the validity of self-report instruments in the field of criminal law (Edens, 2009; Milton et al., 2005) called for a systematic review and meta-analysis of the self-reported levels of antisocial behavior and psychopathic features as well as the maladaptive personality traits of the two most relevant personality dimensions of agreeableness and conscientiousness in forensic populations, i.e., anger, aggression, hostility, and impulsivity.

The study – to the author's knowledge the first comprehensive meta-analytic review of self-reported maladaptive personality characteristics carried out in forensic populations and compared to non-

forensic norm or reference groups – found no overall differences in self-reported levels of anger, aggression, hostility, or impulsivity between the general or healthy population and forensic samples, while self-reported antisocial and psychopathic features were significantly and substantially higher in forensic samples than in reference groups. The latter is in line with other studies that found that forensic subjects with antisocial tendencies or psychopathic features respond truthfully about their levels of these traits (Cima et al., 2008; Edens, 2009; Niesten, Nentjes, Merckelbach, & Bernstein, 2015; Ray et al., 2013).

At the same time, other studies suggest that self-report measures potentially yield biased results, with both under-reporting and over-reporting of personality traits that are considered negative. Previous findings of low levels of self-reported aggression and hostility (Hornsveld et al., 2009) and overall self-reported personality pathology in forensic populations (Spaans et al., 2015) suggest that self-report assessment within forensic samples may underestimate the actual levels of these maladaptive personality traits. Blackburn and colleagues (2004), and Milton and colleagues (2005) also suggest that individuals with antisocial traits are inclined to deny or minimize negative traits, especially when they may have legal repercussions. On the other hand, other studies found that Cluster B personality disorders and their symptoms are underestimated by self-report instruments (Blackburn et al., 2004; De Ruiter & Greeven, 2000; Zimmerman & Coryell, 1990).

The extent to which issues such as malingering (i.e., the exaggeration of negative qualities) or dissimulation and positive impression management (i.e., giving socially desirable answers on items that clearly describe a negative trait) play a role in the way the individual presents him or herself within the forensic mental health context, leading to lower than expected levels of anger, aggression, hostility, and impulsivity in forensic samples, remains unclear. It could also be the case that on some traits, individuals display unintentional self-deception (Ray et al., 2013) or that deceptive or manipulative behaviors are displayed because they are characteristic of Cluster B personality disorder symptoms (Cima, 2003) which are highly prevalent in forensic samples (De Ruiter & Greeven, 2000; Hildebrand & De Ruiter, 2004; Timmerman & Emmelkamp, 2001). In any case, the findings of the current and previous studies indicate a need for caution when using self-report assessment instruments in forensic populations, especially those that do not include a response bias scale.

Another issue that was brought to light in the systematic review and meta-analysis of a number of self-reported maladaptive personality traits was the overwhelming multitude of assessment instruments and consistently differently named subscales. The majority of the instruments used in a forensic mental health setting are not designed especially for forensic populations (Wygant & Lareau, 2015). It is imperative to find a universal instrument and use universal terminology in personality traits and (sub)scale names, that is most suitable for the forensic mental health field. Only then can results be properly compared worldwide. The study also showed that different instruments produced differing levels of self-reported maladaptive personality traits, adding to existing doubts about the validity of self-report instruments in the forensic mental health setting and calling for further research. The actual levels of maladaptive personality traits in forensic samples need to be determined, using alternatives for self-report such as clinician-administered (semi-)structured interviews, observer-rated assessment methods, or cognitive tasks, as well as whether different kinds of deviant response styles apply to different personality traits or different (legal) consequences.

## Chapter Three: The utility of the DAPP-SF as a screener for personality disorder

For various reasons it is important that the forensic assessment process is efficient and cost-effective (Wygant & Lareau, 2015). The two-step approach for efficient assessment of personality psychopathology, recommended by Widiger and Samuel (2005) in the field of general psychological practice, involves first administering a self-report questionnaire to screen for the potential presence of personality disorders, and to follow it with a standardized (semi-)structured diagnostic interview if there are indications of a disorder to verify its presence. If such a screening questionnaire were quick and accurate, decreasing the number of patients who do not have a personality disorder still having to undergo a standardized (semi-)structured diagnostic interview in the field of criminal law, the diagnostic process would become much more efficient and cost-effective.

Given the success of the DAPP-SF as a screener for personality disorder in the general population (De Beurs et al., 2010), the suitability of the DAPP-SF as a screener for personality disorder was investigated in a forensic psychiatric hospital that carries out pre-trial evaluations on suspects of serious crimes. To the author's knowledge, this was the first time this was studied in a forensic population. It was expected that when scores on (sub)scales of the DAPP-SF were elevated, chances of criteria for personality disorder(s) on the Structured Interview for DSM-IV Personality (SIDP-IV) being met were high. Results showed only a few small to medium effect sizes that gave some support for the use of the DAPP-SF as a screener for personality disorders. The main finding, however, was that the forensic population reported significantly less personality pathology than the general population on 14 out of the 22 personality dimensions and second-order factors of the DAPP-SF. Based on this floor effect as well as only moderate associations between DAPP-SF and SIDP-IV outcome it was concluded that the DAPP-SF has limited utility as a screening tool for personality disorders in a forensic context.

Again, the question arises whether the levels of self-reported personality pathology found in the current study reflect the actual levels of the respondents or whether the participants dissimulated or displayed positive impression management when they filled in the self-report questionnaire to decrease their chances of undesirable legal consequences, such as enforced treatment in a high security hospital in the Netherlands. In these cases where the stakes for the respondent are high, self-report instruments are prone to bias and of limited utility. Other methods of psychological assessment, such as expert-based judgments aided by checklists such as the PCL-R, will be superior to self-reports under these circumstances.

## Chapter Four: A cluster analysis of MMPI-2 profiles

Another self-report instrument of particular interest, especially after the systematic review and meta-analysis of self-reported personality traits in which the instrument revealed particularly high levels of antisocial and psychopathic features, is the MMPI. Previous research aimed at classifying criminal offenders according to their MMPI profiles has found between two and ten distinct profile types. Specific information on each profile type of offenders could lead to their tailored treatment and management. As studies resulting in only two profile types raise doubts about the suitability of the MMPI-2 as the

primary instrument for differentiating personality types in a forensic population, the cluster analysis was replicated in a sample of pre-trial criminal defendants in a psychiatric observation hospital.

The current study again found only two groups of distinct MMPI-2 personality profiles: one indicating the absence of any psychological problems and the other exhibiting elevations on all but three scales. These findings were surprising, as the study population was considered diverse – containing not only severe and varied psychopathology but also individuals with no psychological disorders and different types of offenders. It was concluded that the profile types were not effectively distinguished by the MMPI-2. Once again, results raised doubts about the use of a self-report instrument in forensic populations, and the usefulness of interpreting MMPI-2 results for diagnostic purposes in particular. Results suggest that the usefulness of the MMPI-2 in a forensic context is restricted to screening for the presence or absence of general psychopathological symptoms, and that it is not able to distinguish between subtypes of psychopathology.

While the MMPI is popular and widely used in Dutch forensic practice, Walters (2006) argues that it has weak overall content validity and the MMPI's Psychopathic deviate Pd scale is considered very heterogeneous and multifactorial (Lilienfeld, 1999; Osberg & Poland, 2001; Derksen, De Mey, Sloore, & Hellenbosch, 2006). Lally (2003) is also critical of the MMPI, stating that although a test may be popular in a particular field, it should not necessarily be used in that field. He states that the MMPI-2 only provides information about an individual's current response style and psychopathology, not on past features that are relevant to the legal issue, and that it might actually be used in a way that courts might find neither relevant nor reliable.

On the other hand, the findings of the two clusters could also be explained as an egosyntonic profile versus an egodystonic profile. Individuals in the former profile type, corresponding to the cluster with only an elevation on the Pd scale have very little to no insight into their own personality (pathology), while the latter profile, corresponding to the many elevated scales, contains individuals who do have a realistic self-image. Further inspection of the MMPI-2 validity scales of the study sample showed no elevated scores in the non-disturbed cluster, while the F (Low Frequency) scale was elevated in the disturbed cluster. This indicates that the disturbed cluster was not only realistic in their self-image, but could in actual fact have been exaggerating their symptoms. However, in the Dutch legal context with the possibility of enforced treatment without a clear release date, this is not behavior that is often seen. These conflicting interpretations strengthen suggestions that while the MMPI-2 can certainly be used to investigate the presence of overall self-reported psychopathology, it is less suitable for specific diagnostic purposes.

Despite these findings, the MMPI-2(-RF) remains very popular with forensic mental health experts. This is at least in the Netherlands, where forensic mental health experts often feel that there is no suitable or concise alternative self-report instrument for overall personality pathology. It has been suggested that theory-driven combinations of single MMPI scales (instead of single scale scores) may constitute a measure of level of personality organization in theoretically predicted ways (Eurelings-Bontekoe, Luyten, Remijnsen, & Koelen, 2010), but the diagnostic utility of such a profile analysis of the MMPI for the forensic setting awaits further empirical study. This issue stresses the need for an assessment battery that has been specifically designed for the forensic field as well as the need for more research into the value and validity of more complex interpretations of test-results, beyond single subscale scores as suggested by Eurelings-Bontekoe and colleagues (2010). The consequences of the results of a forensic assessment can obviously be quite life changing, much more so than in other areas of psychological assessment (such as employee testing and selection). A thorough and accurate assessment and diagnosis is of the utmost importance in the forensic field.

## Self-report instruments in a forensic mental health setting

The results of these three chapters suggest that although self-report instruments such as the DAPP-SF and the MMPI-2 can be of certain value in individual use and case finding of possible personality pathology, caution should be used when interpreting results. Self-report assessment instruments are generally not designed specifically for the forensic population (Wygant & Lareau, 2015) and should never be the sole source of diagnostic information in forensic assessments, as stipulated in the American Psychological Association's guidelines for forensic psychology (American Psychological Association, 2013b). In order to diagnose personality disorders in forensic participants, self-report assessments should be incorporated into information gathered in other ways, such as through (semi-) structured clinical interviews, extensive collateral information, file information and hetero-anamnestic data (Cima, 2003; Hildebrand & De Ruiter, 2004; Wygant & Lareau, 2015).

Forensic mental health experts should always be aware of the possibility of socially desirable response tendencies. This tendency to give positive self-descriptions (Paulhus, 2002) includes both intentional positive impression management and faking good (presenting oneself in a positive light) and unintentional self-deception (Ray et al., 2013) on items that clearly describe a negative trait. Intentional dissimulators could, for example, be motivated by the preference of a defined prison term over an undefined term of hospitalization and the wish to avoid stigmatization. Uninsightful dissimulation is not a rational choice, but a genuine lack of knowledge or awareness of one's psychiatric disorder or symptoms that also leads to low self-reported personality pathology (Caruso et al., 2003). Gutheil (2003) suggested that inmates might also be encouraged by their attorneys to present their symptoms in a certain, tactical way. This has also been found to be the case in Dutch legal practice (Barendregt, 2010). In any case, as Milton and colleagues (2005) suggested, high security patient samples are "essentially a highly abnormal group of mentally-disordered offenders in a highly unusual hospital setting who may consciously or unconsciously fake good to present themselves in the best possible light to affect their progress or discharge" (p. 559).

## Positive impression management in forensic mental health settings

Proof of dissimulation by positive impression management in a forensic mental health context has been found by several authors (Ahlmeyer et al., 2000; Gutheil, 2003; McEwan et al., 2009; Mills et al., 2003). A number of other authors advise against self-report instruments in forensic populations, unless they contain a measure for positive impression management or dissimulation (De Beurs & Barendregt, 2008; De Ruiter & Greeven, 2000; Edens & Ruiz, 2006; Mills et al., 2003). On the other hand, a meta-analysis carried out by Ray and colleagues (2013) on the response validity scales of self-report assessment instruments of psychopathic traits – the MMPI, the Psychopathic Personality Inventory (PPI), and the Levenson's Self-report Psychopath (LSRP) – found no association between psychopathy and measures of social desirability or faking good.

## Response tendencies and legal consequences

Research by Niesten and colleagues (2014) showed that both faking good and bad are more common in prison settings than in forensic psychiatric settings. In a study using the MMPI on a sample of prison inmates within varying incentives to either refute or exaggerate personality pathology, Walters (1988) found that response styles were related to possible results, such as placement in a single cell, entering group therapy, and applying for parole.

Cima and colleagues (2007) use the term “supernormality” to describe the tendency to systematically deny the presence of common, possible negative symptoms. In an investigation into different forms of deception such as malingering and supernormality, they found that response styles of forensic populations are related to both possible (legal) consequences and personality traits such as psychopathy. Cima and colleagues (2003) suggest that once convicted, offenders would engage in faking good to acquire privileges such as parole. Nentjes, Bernstein, Arntz, Slaats, and Hannemann (2015) found that offenders with good understanding of the perspective of others along with high levels of impulsivity, emotional instability, or aggression display *less* socially desirable traits and in fact report uncharacteristic and unusual psychiatric symptoms. In a study on self-reported psychopathic traits on the PPI-R in male offenders, Kelsey, Rogers, and Robinson (2015) found that the groups with high and moderate psychopathy scores were equally able to display positive impression management and mask their psychopathy. These three studies reinforce the suggestion that response styles in a forensic mental health setting are related to a combination of incentives and personality traits.

Even Ray and colleagues (2013), who found no association between psychopathy and social desirability or faking good, added the caveat to their findings that they were based on guarantees of confidentiality and/or anonymity and that there were no obvious incentives for deviant response tendencies. Their findings do not rule out the possibility that psychopathic individuals do engage in faking good or bad when there are actual incentives to doing so.

It would be interesting to investigate whether similar results for the MMPI-2 and DAPP-SF are found when it is administered to patients already undergoing enforced treatment in high-security hospitals. This would clarify whether it is indeed the pre-trial nature of the study contexts that causes the limited ability of the DAPP-SF to screen for personality disorders and the MMPI-2 to classify offender types, or whether it is the forensic mental health setting in general.

## Chapter Five: Personality disorder and criminal responsibility

In the Netherlands, once personality pathology has been established in pre-trial psychological assessments, judges have to decide on the level of criminal responsibility and on the necessity of enforced treatment in a high security hospital. The first study of its kind, to the author’s knowledge, of the Dutch forensic context consisted of two investigations into whether personality disorders and psychopathic traits in criminal suspects are reasons for diminished criminal responsibility or enforced treatment in high security hospitals. Results demonstrated an internationally unique role of personality disorder compared to other jurisdictions where the presence of (antisocial) personality disorder generally does not lead to diminished criminal responsibility. In the Netherlands, the presence of a personality disorder decreased responsibility and led to an advice for enforced treatment. Results also showed that PCL-R items concerning impulsivity and (ir)responsibility were considered to impair

an individual's freedom of will. High PCL-R Facet 2 and Facet 3 scores were related to an advice for enforced treatment. The results of this study show that Dutch forensic mental health clinicians take the diagnosis of a personality disorder or a high PCL-R score into careful consideration when making recommendations for diminished responsibility or the need for enforced treatment, stressing the importance of a thorough diagnosing of personality pathology in psychological and psychiatric assessments of suspects in criminal cases.

## Chapter Six: Personality disorder and enforced treatment

Once a Dutch court has established a mental disorder at the time of the crime and diminished criminally responsibility, the individual can be sentenced to enforced treatment in a high security hospital. As concluded from the previous chapter, in Dutch forensic practice this mental disorder includes the presence of a personality disorder. In a study to prospectively investigate the predictive value of personality disorder on treatment length, a study sample of 536 forensic psychiatric patients from various high security hospitals with a 76.5% prevalence of personality disorder was investigated. Data from two different judicial phases were merged (from pre-trial assessment and treatment in a high security hospital) and Axis I disorder, substance abuse history, and intellectual functioning were also taken into account as possible confounders.

Results showed no independent predictive value of personality disorder for treatment length at first accompanied leave request and total treatment length at the end of treatment. An index offense of (attempted) violence and (attempted) murder/manslaughter both independently predicted shorter treatment length at first accompanied leave request across different strata for treatment hospitals. An index offense of (attempted) sex offense independently predicted a longer total enforced treatment length, again across different strata for treatment hospitals. A post-hoc explorative analysis on a subgroup of patients with a diagnosis of *only* personality disorder showed that this group had a shorter treatment length than personality disordered patients who also had a comorbid Axis I disorder. Analysis of the influence of a proxy for time spent in prison before treatment on total treatment length showed a negative relationship for patients with a diagnosis of *only* personality disorder. This might indicate that a personality disorder may worsen during imprisonment prior to treatment and subsequently take longer to treat, although it must be noted that this was an explorative analysis carried out on a small subset of patients.

This study also showed that investigation of the factors involved in the prediction of length of enforced treatment of disordered offenders can be complex, due to limited (digital) data registration and challenges in combining data from separate databases. This study can be considered a first attempt at distilling these factors from available rough, unrefined digital data and calls for better compatibility of data registration and collecting and an increase in (detail of) digital registration of patient, treatment, and risk characteristics along the course of treatment and at its end.

## Enforced treatment of personality pathology

The enforced treatment of (personality-)disordered offenders is a complex issue, and according to Van der Wolf (2012) will never be smooth sailing. It is clear that serious personality pathology is very

difficult to change (Hildebrand & De Ruiter, 2012), especially Cluster B personality disorders (Bernstein et al., 2012). Types of treatments for personality disordered offenders range from behavioral (targeting observable actions) to psychoanalytical (targeting abstract mental representations) but generally incorporate elements from both methods into one approach. This can include psycho-education and the development of skills such as reflection and self-awareness (NOMS, 2011). In a review of available treatments for severe personality disorder, Warren and colleagues (2003) concluded that long-term treatment in a therapeutic community institution was the most effective in reducing recidivism when compared to untreated inmates who remained in the general prison system.

According to Hornsveld (2007), treatment programs for forensic psychiatric inpatients should focus on improvement of anger management, social skills and antisocial attitudes. De Ruiter, Veen, and Greeven (2008) performed a meta-analytic review of psychological interventions for adult rapists and found only modest effects. They concluded that the most effective interventions were intensive, lengthy and inpatient programs. Van den Berg and Oei (2009) claim that when treating personality disorders, it is of great importance that the many different available therapeutic approaches are all exploited in order to effectively cover the multifaceted complexity of the disorder. They encourage 'social treatment' of patients with antisocial personality disorder, while dealing with patients and while collaborating with other clinicians. They also suggest that cognitive behavioral therapy, group therapy, and Mentalization Based Treatment can strengthen each other in the treatment of forensic psychiatric patients with personality disorders. According to the National Offender Management Service, successful treatments combine group and individual treatment, and include a team approach and intensive contact with the patient (NOMS, 2011).

Bernstein and colleagues (2012) studied Schema Therapy (ST), which combines cognitive, behavioral, psychodynamic object relations, and humanistic or experimental approaches and was developed for patients with persistent cognitive, emotional, and behavioral patterns who profited insufficiently from cognitive behavioral therapy or classic forms of psychotherapy (Kool & Aalders, 2014). In a study in a population of forensic psychiatric patients with antisocial, borderline, narcissistic and/or paranoid personality disorders in Dutch high security hospital, Bernstein and colleagues (2012) found that ST does not yield better outcomes than treatment as usual with regard to recidivism risk and return to society. A planned three-year follow-up study will examine the long-term effects of ST on actual recidivism of these patients.

As with personality disordered offenders, the treatment of psychopathic offenders is also challenging. This is because the disorder is quite complex and psychopathic offenders display disruptive behavior during treatment, are likely to drop out, display low levels of change (Hildebrand & De Ruiter, 2012), and might even be adversely impacted by treatment and display higher recidivism (Harris, Rice, & Cormier, 1994; Hildebrand & De Ruiter, 2012). Still, this does not mean that there is no hope in treating psychopathic offenders. According to Hare (2006), although traditional programs have not shown much effect, innovative procedures designed specifically for psychopathic offenders need to be developed. After meta-analytic research of on the treatment of psychopathy, Salekin, Worley, and Grimes (2010) concluded that treatment shows moderate success. They recommend that researchers clearly map out all the areas and processes of dysfunction in psychopaths, followed by clear theories for therapeutic change, possibly in controlled studies.

## Faking good during treatment

As with diagnosing personality pathology, its treatment can also be subject to different kinds of deviant response styles. Socially desirable response tendencies, faking good, positive impression management and supernormality may lead to an erroneous impression that progress is being made when, in fact, the patient still has psychopathological or psychopathic symptoms. Nijman and colleagues (2002) found that while personality disordered criminals showed significantly more treatment progress than psychotic forensic psychiatric patients, improvement during treatment does not necessarily imply that the risk of recidivism after release has decreased. Indications of progress during treatment rely heavily on self-reports of patients. Personality disordered patients, especially antisocial or psychopathic patients, might maintain or develop a tendency to endorse non-symptomatic answer options (Cima et al., 2003). Ray and colleagues (2013) concur, stating that because there is no objective criterion or golden standard for accurate responding on self-report measures of psychopathic traits, the possibility that psychopathic forensic psychiatric patients with low scores on social desirability measures are nonetheless underreporting negative attributes cannot be excluded.

## Conclusion

In summary, it can be said that, although personality pathology such as maladaptive personality traits are highly prevalent in forensic mental health settings according to various studies, chapters Two, Three and Four of this dissertation have shown that personality pathology of forensic populations is difficult to assess through self-report questionnaires. This is due to the possibility of deceptive or manipulative response tendencies when individuals are asked to present themselves through self-report assessment instruments, as well as unintentional self-deception due to lack of self-insight. Self-report instruments can be of certain value in individual use and case finding of possible personality pathology, but caution should be used when interpreting results. Response bias scales should always be included and results interpreted with great caution or even dismissed when elevated scores on these bias indicators are attained, or when regular scales scores are low and the forensic mental health expert suspects a response bias. Self-report instruments should never be the sole information source in forensic assessments, and should be incorporated into information gathered in other ways, such as through (semi-)structured clinical interviews, observation, cognitive tasks, extensive collateral information, file information and hetero-anamnestic data.

Chapters Five and Six have shown that, in the Dutch forensic context, personality disorder is seen as possible reason for diminished criminal responsibility and enforced treatment, and indicated that a small subset of individuals who suffer from only personality disorder have a shorter enforced treatment length in comparison to personality disordered patients who also had a comorbid Axis I disorder. This first attempt at distilling the factors involved in the prediction of a successful resocialization process of disordered offenders, from limited sources of digital data and differing databases per judicial organization, has isolated separate factors that can shorten or prolong treatment length and can provide a focus for both policy and (clinical) decision makers in the enforced treatment process (e.g., an index offense of (attempted) violence and (attempted) murder/manslaughter for a shorter treatment length and an index offense of (attempted) sex offense for a longer treatment length).

One of the ways to maximize treatment effectiveness for offenders with personality pathology is through an increase in (detail of) digital registration of patient, treatment process, outcome, and risk characteristics along the course of and end of treatment.

## Study limitations

A limitation of the studies presented in this dissertation is their generalizability outside the Dutch context. All but the first study were carried out within the Dutch forensic and legal context in which, as shown in Chapter Five, personality disorder plays an internationally unique role. It should be noted that the Dutch legal system of deciding criminal responsibility differs substantially from other jurisdictions (Brants, 2008), and that therefore the results in these studies are difficult to generalize to the U.S., Canada, or other European countries. Limitations to the generalizability of the findings also present themselves within the Dutch context. The study populations used in Chapters Three and Five included only cooperating, sufficiently Dutch-speaking, nonpsychotic participants who had been admitted to an observation hospital, leading to the theoretical possibility of selection bias and reservations about the generalizability of the results to outpatients, convicted offenders, or those suspected or convicted of less severe crimes than the PBC's population. The study presented in Chapter Four included only males. The study presented in Chapter Six included defendants who had been evaluated in either an in-patient (observation hospital) or an out-patient (regular detention) setting. It must be noted that the system of enforced treatment in the Netherlands has undergone so many frequent and rapid policy changes over the past years that this may have influenced the results of Chapter Six.

A limitation of the meta-analysis presented in Chapter Two is that the formulation of the three overarching categories of maladaptive personality traits could have been approached differently, as some personality traits may share features of more than one domain. If certain personality characteristics had been placed in a different category, results may have been different. A limitation of the examination of the utility of the DAPP-SF as a screener for personality disorder, presented in Chapter Three, is that it could not be formally assessed whether the suggested positive impression management was actually displayed by the study population. An alternative explanation for the findings could be lack of statistical power due to the small sample size. A further limitation was the lack of data on interrater reliability for the PBC's forensic mental health experts on the SIDP-IV. However, a 72.5% convergence between SIDP-IV classifications and clinical diagnoses, along with the fact that most dissimilarities were extensively clarified by the experts in their final reports, supported the decision to use the SIDP-IV as the criterion for the presence of personality disorder.

The use of data from diverse digital databases, in the study on the prognostic value of personality disorder on enforced treatment length presented in Chapter Six, introduced a few additional limitations. It was not possible to incorporate patient characteristics from the actual treatment (such as type of treatment, treatment process, or treatment outcome) as this data was not digitally available. In some cases more than ten (treatment) years had passed since the NIFP's diagnosis. Besides the possibility of human error occurring when manually entering data, researchers are dependent on the level of detail that has been chosen for each digital database, which was relatively low in the present case. Therefore, it was not possible to investigate aspects such as possible comorbidity within personality disorders or differences in severity of (personality) disorders.

## Implications of the results for clinical practice

The main implication of the results of the current dissertation is to proceed with caution at all times when using self-report assessment instruments for personality pathology in forensic populations. This is especially the case when those self-report instruments are not designed specifically for the forensic population and do not include a response bias scale. The studies in Chapters Two, Three, and Four mean that forensic mental health experts should at all times be aware that results of self-report assessment instruments are prone to bias due to intentional impression management or unintentional self-deception. Findings of forensic assessments should be based on a combination of assessment methods: (semi-)structured clinical interviews, extensive collateral information, cognitive tasks, file information and hetero-anamnestic data. Self-report instruments can still be of certain value in individual use and case finding of possible personality pathology but should always be incorporated into information gathered through the above-mentioned other ways.

Use of the DAPP-SF and the MMPI-2 in particular, for specific diagnostic purposes in forensic populations, should be accompanied with caution. The current study found that the DAPP-SF has limited usefulness as a screener for personality disorders and that profile types are not effectively distinguished by the MMPI-2.

Also, there should be more standardization of terminology and assessment instruments concerning personality pathology in the forensic mental health field so that study outcomes can be more easily compared with each other. As mentioned above, the most effective treatment methods for personality disordered offenders appear to be those that are long-term and intensive, and possibly should incorporate several different therapeutic approaches so that they can strengthen each other. Furthermore, there should be an increase in (detail of) digital registration of patient, treatment, and risk characteristics along the course of and end of treatment, in order to facilitate research on treatment effectiveness for offenders with personality pathology.

One of the major hurdles in the Dutch forensic psychiatric field is the fear of enforced treatment among defendants accused of crimes that are sufficiently serious to potentially warrant enforced treatment. They generally perceive enforced treatment as far more taxing than a prison sentence, largely due to the uncertainty about the length of treatment. Some argue that the length of time spent in high security treatment hospital is often longer than the amount of time the individual would have spent in prison for the same offense. It is also said that the stigma of a forensic psychiatric patient as a “dangerous lunatic” reduces the chances of successful resocialization (Van der Wolf, 2012). Steps have been suggested to reduce the fear of enforced treatment and thus increase the levels of participation in pre-trial psychological and psychiatric assessments and the data that would become available on a wider variety of defendants and personality pathology. These suggestions include to place more emphasis on quicker resocialization of forensic psychiatric patients to increase the national confidence in enforced treatment, to offer a “trial-treatment” to demonstrate that it might not be as bad as it sounds, to impose an additional prison sentence in cases of refusal to cooperate with the pre-trial assessment, or to reward cooperation, financially or otherwise (Van der Wolf, 2012).

## Suggestions for further research

A thorough and accurate diagnosis of personality pathology is of the utmost importance in the forensic mental health field. The consequences of the results of a forensic pre-trial assessment can be much more life changing than in other areas of psychological assessment. There is an urgent need for an assessment battery that has been specifically designed for the forensic mental health field, and for a concise screening instrument for personality disorders to successfully identify individuals that require further assessment. It is also essential to introduce more standardization of terminology and assessment instruments for personality pathology to be able to properly compare forensic results worldwide. A first step towards this standardization would be to no longer use assessment instruments that have not been validated for forensic populations or designed specifically for them. This would already eliminate a great number of instruments.

Implications for future research in forensic populations also include a detailed investigation into response bias tendencies within differing legal contexts, such as convicted offenders instead of suspects or those undergoing enforced treatment, to examine the possibility of legal incentive-related response bias tendencies. This would hopefully clarify whether the pre-trial setting caused the limited ability of the DAPP-SF to screen for personality disorders and the MMPI-2 to classify offender types, or whether it is the forensic mental health setting in general that leads to these results.

In order to measure levels of maladaptive personality traits in forensic samples in a valid way, alternatives for self-report such as clinician-administered (semi-)structured interviews, observer-rated assessment methods, or cognitive tasks should be used. Also, an investigation into whether different kinds of deviant response styles apply to different personality traits or different (legal) consequences is necessary.

As mentioned above, treatment methods for personality disordered offenders should be long-term, intensive, and possibly incorporate several different therapeutic approaches. The evidence-base for the best approach, however, remains limited. To make the most of the long periods of time needed to assess the effectiveness of treatment (most often measured by the rate of recidivism upon release after a certain follow-up period) and not lose opportunities for research, an increase in (detail of) digital registration of patient, treatment, and risk characteristics along the course of treatment and at its end is imperative. As many outcome studies in the past have shown methodological weaknesses (Warren et al., 2003), efforts should be made to avoid these weaknesses. Although this may be quite challenging to realize within forensic mental health settings, future research should preferably consist of randomized-controlled studies, and give detailed information on the study population and selection criteria, the methods of diagnosing the subjects (which should be validated for forensic populations), on the treatment(s) administered (to enable replicating this treatment), and on the choice of (standardized) outcome measure and statistical analyses. Studies should also use large samples and comparison groups, address possible attrition, and separate Axis I and Axis II diagnostic categories where possible. Carrying out further meaningful research on the effectiveness of intensive and elaborate treatment, and investigating which patients benefit most from which treatment at which stage, will hopefully contribute to improved (risk) diagnoses, fewer incidents, shorter treatment duration, and less recidivism.

As mentioned earlier, there should be an increase in detail of digital registration of patient, treatment, and risk characteristics over the course of treatment. In order to facilitate research on assessment, diagnostics, and treatment effectiveness for offenders with personality pathology, legal and psychiatric organizations should work together to create complete and compatible databases that cover the entire legal and psychiatric process from arrest to release of the offender.

De Beurs and Barendregt (2008) investigated suitable ways to study the effect of enforced treatment, meant to facilitate the development of evidence based treatment programs. After taking into consideration important criteria for research design – such as randomization, statistical power, treatment integrity and protocols, willingness of the treatment hospital to apply assessment instruments, as well as logistics and infrastructure – they concluded that due to practical and ethical challenges to randomized-controlled studies in the forensic mental health setting, outcome monitoring with cohort studies is the most suitable method.

Outcome monitoring entails periodically assessing all patients' conditions or symptoms and general psychological wellbeing, using the same assessment instrument each time, and recording what type and quality of treatment each patient is receiving (De Beurs & Barendregt, 2008). In the Netherlands this practice is known as Routine Outcome Monitoring (ROM), which, as mentioned earlier, is aimed at transparency, evaluation and possible adjustment of treatment. The authors also stress the importance of organizations such as treatment hospitals working together to combine sufficient data concerning the effects of specific treatment programs. This first requires that all organizations use the same assessment instruments and record data in exactly the same way.

## Final remarks

The main conclusion of this current dissertation is that personality pathology, with its high prevalence in the forensic mental health setting, plays a pivotal role in the assessment, diagnosis, treatment and return to society of disordered offenders. It remains under-researched, however, and very few assessment instruments have been designed specifically for this population. Available assessment instruments that are known to be reliable and valid in the general population should be used with caution among defendants or offenders. The importance of sound and thorough assessment, diagnosis, and treatment of offenders – both for the offender and for society – calls for assessment instruments specifically designed for this population and an increase in (detail of) digital registration of patient, treatment, and risk characteristics.

# References

References marked with an asterisk indicate studies included in the meta-analysis.

- \*Adams, T. C. (1976). Some MMPI differences between first and multiple admissions within a state prison population. *Journal of Clinical Psychology*, 32(3), 555-558.
- Ahlmeyer, S., Heil, P., McKee, B., & English, K. (2000). The impact of polygraphy on admissions of victims and offenses in adult sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 12(2), 123-138.
- Almagor, M. & Nevo, B. (1996). The MMPI-2: Translation and first steps in its adaptation into Hebrew. In: J. N. Butcher (Ed), *International adaptations of the MMPI-2: Research and clinical applications* (pp. 487-505). Minneapolis, MN: University of Minnesota Press.
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- American Psychiatric Association (2013a). *Diagnostic and Statistical Manual of mental disorders: DSM-5*. Washington, D.C.: American Psychiatric Association.
- American Psychiatric Association (2013b). Specialty guidelines for forensic psychology. *American Psychologist*, 68(1), 7-19.
- Anderson, J. L., Sellbom, M., Wygant, D. B., & Edens, J. F. (2013). Examining the necessity for and utility of the Psychopathic Personality Inventory-Revised (PPI-R) validity scales. *Law and Human Behavior*, 37(5), 312-320.
- Arnett, J. (1994). Sensation seeking: A new conceptualization and a new scale. *Personality and Individual Differences*, 16, 289-296.
- AVT (2011). *Nieuwsbrief nr. 6* [Newsletter no. 6]. Utrecht: AVT.
- AVT (2014). *Jaarverslag 2013* [Annual report 2013]. Utrecht: AVT.
- Baillie, A. J., & Lampe, L. A. (1998). Avoidant personality disorder: Empirical support for the DSM-IV revisions. *Journal of Personality Disorders*, 12, 23-30.
- Barelds, D. P. H., Luteijn, F., Van Dijk, H., & Starren, H. (2007). *Handleiding Nederlandse Persoonlijkheidsvragenlijst 2 (NPV-2)* [Manual of the Dutch Personality Questionnaire]. Amsterdam: Harcourt Publishers.
- Barendregt, M. (2006). Strafrecht, psychologie en wetenschap: over Mooij's Toerekeningsvatbaarheid [Criminal law, psychology and science: About Mooij's criminal responsibility]. *Nederlands tijdschrift voor de psychologie*, 61, 31-38.
- Barendregt, M. (2010). *Onderzoek toenemend aantal weigeraars pro Justitia rapportage en afnemend aantal tbs-opleggingen* [Research on increasing number of non-cooperating defendants of pre-trial evaluations and decreasing number of court-ordered forensic treatments]. Utrecht: NIFP.
- Barendregt, M., Muller, E., Nijman, H., & De Beurs, E. (2008). Factors associated with experts' opinions regarding criminal responsibility in The Netherlands. *Behavioral Sciences and the Law*, 26, 619-631.
- Barendregt, M., Rijnders, R., & Ranitz, A. (2008). Rapporteren bij ontkennende verdachten? Een reactie op De Ruiters [Reporting on suspects who deny their crime? A reaction to De Ruiters]. *Proces*, 1, 20-26.

- Barratt, E. S. (1985). Impulsiveness subtraits: Arousal and information processing. In: J. T. Spence, C. E. Itard (Eds.), *Motivation, emotion, and personality* (pp. 137-146). Amsterdam: Elsevier.
- \*Bauer, G. E., & Clark, J. A. (1976). Personality deviancy and prison incarceration. *Journal of Clinical Psychology*, 32(2), 279-283.
- Berghuis, H. (2014). *General personality disorder: A study into the core components of personality pathology*. Amsterdam: Universiteit van Amsterdam.
- Bernstein, D. P., Iscan, C., & Maser, J. (2007). Opinions of personality disorder experts regarding the DSM-IV personality disorders classification system. *Journal of Personality Disorders*, 21, 536-551.
- Bernstein, D. P., Nijman, H. L. I., Karos, K., Keulen-De Vos, M., De Vogel, V., & Lucker, T. P. (2012). Schema therapy for forensic patients with personality disorders: Design and preliminary findings of a multicenter randomized clinical trial in the Netherlands. *International Journal of Forensic Mental Health*, 11, 312-321.
- Black, D. W., Gunter, T., Loveless, P., Allen, J., & Sieleni, B. (2010). Antisocial personality disorder in incarcerated offenders: Psychiatric comorbidity and quality of life. *Annals of Clinical Psychiatry*, 22(2), 113-120.
- Blackburn, R., Donnelly, J. P., Logan, C., & Renwick, S. J. D. (2004). Convergent and discriminate validity of interview and questionnaire measures of personality disorder in mentally disordered offenders: A multitrait-multimethod analysis using confirmatory factor analysis. *Journal of Personality Disorders*, 18(2), 129-150.
- \*Blagov, P. S., Patrick, C. J., Lilienfeld, S. O., Powers, A. D., Phifer, J. E., Venables, N., ... Cooper, G. (2011). Personality constellations in incarcerated psychopathic men. *Personality Disorders: Theory, Research, and Treatment*, 2(4), 293-315.
- Bleichrodt, N., & Van den Berg, R. H. (1999). *Handleiding Multiculturele Capaciteiten Test - Middelbaar niveau*. MCT-M [Manual Multi-cultural Capacity Test – Medium level]. Amsterdam: Stichting NOA.
- \*Boccaccini, M. T., Murrie, D. C., Hawes, S. W., Simpler, A., & Johnson, J. (2010). Predicting recidivism with the Personality Assessment Inventory in a sample of sex offenders screened for civil commitment as sexually violent predators. *Psychological Assessment*, 22(1), 142-148.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2005). *Comprehensive Meta-Analysis, Version 2*. Englewood, NJ: Biostat.
- Brand, E. F. J. M., Van Emmerik, J. L., & Raes, B. C. M. (2009). TBS: Van de nood een deugd maken [TBS: Making a virtue of necessity]. *Sancties*, 1, 38-47.
- Brand, E. F. J. M., & Van Gemmert, A. A. (2009). *Toenemende verblijfsduur in de tbs: De ontwikkeling van de gemiddelde duur in de laatste 20 jaar* [Increasing length of stay in the TBS: The developments of the average length in the last 20 years]. The Hague: DJI-CU-AID.
- Brants, C. (2008). Strafrechtsvergelijking [Comparison of criminal law]. *Delikt & Delinkwent*, 3, 214-242.
- Buss, A. H. & Durkee, A. (1957). An inventory for assessing different kinds of hostility. *Journal of Consulting Psychology*, 21(4), 343-349.
- Buss, A. H. & Perry, M. P. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology*, 63, 452-459.
- Buss, A. H. & Plomin, R. (1975). *A temperament theory of personality development*. New York: John Wiley.
- Butcher, J. N., Hass, G. A., Greene, R. L., & Nelson, L. D. (2015). *Using the MMPI-2 in forensic assessment*. Washington, DC: American Psychological Association.
- Carlier, I. V. E., Meuldijk, D., Van Vliet, I. M., Van Fenema, E., Van der Wee, N. J. A., & Zitman, F. G. (2010). Routine outcome monitoring and feedback on physical or mental health status: Evidence and theory. *Journal of Evaluation in Clinical Practice*, 18, 104-110.

- Caruso, K. A., Benedek, D. M., Auble, P. M., & Bernet, W. (2003). Concealment of psychopathology in forensic evaluations: A pilot study of intentional and un insightful dissimulators. *Journal of the American Academy of Psychiatry and the Law*, 31, 444-450.
- Cima, M. J. (2003). *Faking good, bad, and ugly: Malingering in forensic psychiatric inpatients*. Maastricht: Universiteit Maastricht.
- Cima, M., Merckelbach, H., Hollnack, S., Butt, C., Kremer, K., Schellbach-Matties, R., & Muris, P. (2003). The other side of malingering: Supernormality. *The Clinical Neuropsychologist*, 17(2), 235-243.
- Cima, M., Pantus, M., & Dams, L. (2007). Simulation und Dissimulation in Abhängigkeit vom strafrechtlichen Kontext und der Persönlichkeit. [Simulation and dissimulation in relation to criminal context and personality]. *Praxis der Rechtspsychologie*, 17(1), 47-62.
- Cima, M., Van Bergen, S., & Kremer, K. (2008). Development of the Supernormality Scale-Revised and its relationship with psychopathy. *Journal of Forensic Sciences*, 53(4), 875-981.
- Ciocchetti, C. (2003). The Responsibility of the Psychopathic Offender. *Philosophy, Psychiatry & Psychology*, 10(2), 175-183.
- \*Claes, L., Tavernier, G., Roose, A., Bijttebier, P., Smith, S. F., & Lilienfeld, S. O. (2014). Identifying personality subtypes based on the Five-Factor Model Dimensions in male prisoners: Implications for psychopathy and criminal offending. *International Journal of Offender Therapy and Comparative Criminology*, 58(1), 41-58.
- Clark, L. A. (1996). *Schedule for Nonadaptive and Adaptive Personality: Manual for administration, scoring, and interpretation*. Minneapolis: University of Minnesota Press.
- Clark, L.A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification of psychopathology: Challenges to the current system and future directions. *Annual Review of Psychology*, 46, 121-153.
- Coolidge, F. L., Segal, D. L., Cahill, B. S., & Archuleta, J. L. (2008). A new five factor model of psychopathology: Preliminary psychometric characteristics of the five-dimensional personality test (5DPT). *Personality and Individual Differences*, 44, 1326-1334.
- Costa, P. T., & McCrae, R. R. (1990). Personality disorders and the five-factor model of personality. *Journal of Personality Disorders*, 4, 362-371.
- Costa, P. T. Jr., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, 81(2), 322-331.
- Crawford, M. J., Koldobsky, N., Mulder, R., & Tyrer, P. (2011). Classifying personality disorder according to severity. *Journal of Personality Disorders*, 25(3), 321-330.
- Cunradi, C. B., Todd, M., Duke, M., & Ames, G. (2009). Problem drinking, unemployment, and intimate partner violence among a sample of construction industry workers and their partners. *Journal of Family Violence*, 24, 63-74.
- \*Cuomo, C., Sarchiapone, M., Di Giannantonio, M., Mancini, M., & Roy, A. (2008). Aggression, impulsivity, personality traits, and childhood trauma of prisoners with substance abuse and addiction. *American Journal of Drug and Alcohol Abuse*, 34(3), 339-345.
- Damen, K. F., De Jong, C. A., & Van der Kroft, P. J. (2004). Interrater reliability of the structured interview for DSM-IV personality in an opioid-dependent patient sample. *European Addiction Research*, 10, 99-104.
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, 10, 85-104.
- De Beurs, E., & Barendregt, M. (2008). *Mogelijkheden voor therapie-effectonderzoek in de tbs-sector: komen tot een evidence base onder zorgprogramma's* [Possibilities for treatment outcome research in the TBS sector: coming to an evidence base for care programs]. Utrecht/The Hague: NIFP/WODC.

- De Beurs, E., Rinne, T., Van Kampen, D., Verheul, R., & Andrea, H. (2009). Reliability and validity of the Dutch Dimensional Assessment of Personality Pathology – Short Form (DAPP-SF), a shortened version of the DAPP-Basic Questionnaire. *Journal of Personality Disorders*, 23(3), 308-326.
- De Beurs, E., Rinne, T., Van Kampen, D., Verheul, R., & Andrea, H. (2010). Criterion-related validity of the DAPP-SF and its utility as a screener for personality disorders in outpatient care. *Personality and Mental Health*, 4, 271-283.
- De Kogel, C. H., & Nagtegaal, M. H. (2006). *Gewelds- en zedendelinquenten met psychische stoornissen: Een inventarisatie van wetgeving en praktijk in het forensisch psychiatrisch systeem en het gevangenisstelsel in Engeland, Duitsland, Canada, Zweden en België* [Serious violent and sexual offenders with a mental disorder: An inventory of legislation and practical experience in the forensic psychiatric systems and the prison systems in England, Germany, Canada, Sweden and Belgium]. The Hague: WODC.
- De Ruiter, C. (2007). Een dubieuze gedragskundige rapportage bij een ontkennende verdachte. De casus van de Anjummer pensionmoorden [A dubious report by a behavioural expert on a suspect who denied the crime. The case of the murders at Anjum guesthouse]. *Proces*, 4, 136-145.
- De Ruiter, C., & Greeven, P. G. J. (2000). Personality disorders in a Dutch forensic psychiatric sample: Convergence of interview and self-report measures. *Journal of Personality Disorders*, 14(2), 162-170.
- De Ruiter, C., Veen, V., & Greeven, P. (2008). De effectiviteit van psychologische behandeling bij volwassen verkrachters: Bevindingen van een meta-analyse [The effectiveness of psychological treatment of adult rapists: Findings of a meta-analysis]. *Tijdschrift voor Seksuologie*, 32, 28-41.
- \*Dear, G. E., Watt, B. D., & Dockerill, J. (2003). Factor structure of the Spielberger Anger Expression Scale when used with Australian prisoners. *Psychological Reports*, 92, 617-620.
- Del Giudice, M., Booth, T., Irwing, P. (2012). The distance between Mars and Venus: Measuring global sex differences in personality. *PLoS ONE*, 7(1), 1-8.
- Derksen, J., De Mey, H., Sloore, H., & Hellenbosch, H. (2006). *MMPI-2 Handleiding voor afname, scoring en interpretatie* [MMPI-2: Manual for testing, scoring and interpretation]. Nijmegen: PEN Tests Publisher.
- Dienst Justitiële Inrichtingen (2011). *Forensische zorg in getal. 2006-2010* [Forensic care in numbers: 2006-2010]. The Hague: DJI.
- Dienst Justitiële Inrichtingen (2014). *Forensische zorg in getal. 2009-2013* [Forensic care in numbers: 2009-2013]. The Hague: DJI.
- Dienst Justitiële Inrichtingen (2015). *Forensische zorg in getal. 2010-2015* [Forensic care in numbers: 2010-2015]. The Hague: DJI.
- Dolan, M., & Blackburn, R. (2006). Interpersonal factors as predictors of disciplinary infractions in incarcerated personality disordered offenders. *Personality and Individual Differences*, 40, 897-907.
- Doyle, H., & Dolan, M. (2006). Evaluating the validity of anger regulation problems, interpersonal style, and disturbed mental state for predicting inpatient violence. *Behavioral Sciences and the Law*, 24, 783-798.
- \*Echeburúa, E., Fernández-Montalvo, J., & Amor, P. J. (2003). Psychopathological profile of men convicted of gender violence: a study in the prisons of Spain. *Journal of Interpersonal Violence*, 18(7), 798-812.
- \*Edens, J. F. (2009). Interpersonal characteristics of male criminal offenders: personality, psychopathological, and behavioral correlates. *Psychological Assessment*, 21(1), 89-98.

- Edens, J. F., Buffington-Vollum, J. K., Colwell, K. W., Johnson, D. W., & Johnson, J. K. (2002). Psychopathy and institutional misbehavior among incarcerated sex offenders: A comparison of the Psychopathy Checklist–Revised and the Personality Assessment Inventory. *International Journal of Forensic Mental Health, 1*, 49–58.
- Edens, J. F., & Ruiz, M. A. (2006). On the validity of validity scales: The importance of defensive responding in the prediction of institutional misconduct. *Psychological Assessment, 18*(2), 220–224.
- Edwards, A. L. (1959). *Edwards Personal Preference Schedule Manual, revised edition*. New York: Psychological Corporation.
- Espelage, D. L., Cauffman, E., Broidy, L., Piquero, A. R., Mazerolle, P., & Steiner, H. (2003). A cluster-analytic investigation of MMPI profiles of serious male and female juvenile offenders. *Journal of the American Academy of Child & Adolescent Psychiatry, 42*, 770–777.
- Eurelings-Bontekoe, E. H. M., Luyten, P., Remijnen, M., & Koelen, J. (2010). The Relationship between personality organization as assessed by theory-driven profiles of the Dutch short form of the MMPI and self-reported features of personality organization. *Journal of Personality Assessment, 92*(6), 599–609.
- Eysenck, S. B. G., Pearson, P. R., Easting, G., & Allsopp, J. F. (1985). Age norms for impulsiveness, venturesomeness and empathy in adults. *Personality and Individual Differences, 6*(5), 613–619.
- Fazel, S., & Danesh, J. (2002). Serious mental disorder in 23 000 prisoners: a systematic review of 62 surveys. *The Lancet, 359*, 545–550.
- \*Fernández-Montalvo, J., Echauri, J. A., Martínez, M., & Azcárate, J. M. (2012). Batterer men in prison and in court-referred treatment programmes: what is the difference? *Spanish Journal of Psychology, 15*(1), 315–322.
- Fine, C., & Kennet, J. (2004). Mental impairment, moral understanding and criminal responsibility: Psychopathy and the purpose of punishment. *International Journal of Law and Psychiatry, 27*, 425–443.
- \*Firestone, P., Bradford, J. M., Greenberg, D. M., & Larose, M. R. (1998). Homicidal sex offenders: Psychological, phallometric, and diagnostic features. *Journal of the American Academy of Psychiatry and the Law, 26*(4), 537–552.
- Galvan, A., Hare, T., Voss, H., Glover, G., & Casey, B. J. (2007). Risk taking and the adolescent brain: Who is at risk? *Developmental Science, 10*, F8–F14.
- Grann, M., Långström, N., Tengström, A., & Stålenheim, E. G. (1998). Reliability of file-based retrospective ratings of psychopathy with the PCL-R. *Journal of Personality Assessment, 70*, 416–426.
- Gutheil, T. G. (2003). Reflections on coaching by attorneys. *Journal of the American Academy of Psychiatry and the Law, 31*, 6–9.
- Guy, L. S., Poythress, N. G., Douglas, K. S., Skeem, J. L., & Edens, J. F. (2008). Correspondence between self-report and interview-based assessments of antisocial personality disorder. *Psychological Assessment, 20*(1), 47–54.
- \*Haden, S. C., & Shiva, A. (2008). Trait impulsivity in a forensic inpatient sample: an evaluation of the Barratt impulsiveness scale. *Behavioral Sciences and the Law, 26*(6), 675–690.
- Hall, G. C. N., Graham, J. R., & Shepherd, J. B. (1991). Three methods of developing MMPI taxonomies of sexual offenders. *Journal of Personality Assessment, 56*, 2–13.
- Hanson, R. K., & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior, 24*, 119–136.
- Hare, R. D. (2003). *Hare Psychopathy Checklist-Revised technical manual* (2nd ed.). Toronto: Multi-Health Systems.

- Hare, R. D. (2006). Psychopathy: A clinical and forensic overview. *Psychiatric Clinics of North America*, 29(3), 709-724.
- Harris, G. T., Rice, M. E., & Cormier, C. A. (1994). Psychopaths: Is a therapeutic community therapeutic? *Therapeutic Communities*, 15, 283-299.
- Hathaway, S. R. & McKinley, J. C. (1942). The Minnesota Multiphasic Personality Schedule. *National Rehabilitation News*, 9, 20-22.
- Hathaway, S. R., & McKinley, J. C. (1989). *Minnesota Multiphasic Personality Inventory-2: Manual for administration and scoring*. Minneapolis, MN: University of Minnesota Press.
- \*Hepper, E. G., Hart, C. M., Meek, R., Cisek, S., & Sedikides, C. (2014). Narcissism and empathy in young offenders and non-offenders. *European Journal of Personality*, 28(2), 201-210.
- Herpertz, S. C. & Sass, H. (2000). Emotional deficiency and psychopathy. *Behavioral Sciences & the Law*, 18, 567-580.
- \*Herpertz, S. C., Werth, U., Lukas, G., Qunaibi, M., Schuerkens, A., Kunert, H. J., ... Sass, H. (2001). Emotion in criminal offenders with psychopathy and borderline personality disorder. *Archives of General Psychiatry*, 58(8), 737-745.
- 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, 233-248.
- Hildebrand, M., & De Ruiter, C. (2012). Psychopathic traits and change on indicators of dynamic risk factors during inpatient forensic psychiatric treatment. *International Journal of Law and Psychiatry*, 35, 276-288.
- \*Holland, T. R., & Holt, N. (1975). Personality patterns among short-term prisoners undergoing presentence evaluations. *Psychological Reports*, 37(3), 827-836.
- Hornsveld, R. H. J. (2007). *Assessment and treatment of violent forensic psychiatric patients with a conduct or an antisocial personality disorder in the Netherlands*. Nijmegen: Radboud Universiteit Nijmegen.
- Hornsveld, R. H. J., Muris, P., Kraaimaat, F. W., & Meesters, C. (2009). Psychometric properties of the Aggression Questionnaire in Dutch violent forensic psychiatric patients and secondary vocational students. *Assessment*, 16, 181-192.
- Huberty, J. C., DiStefano, C., & Kamphaus, R. W. (1997). Behavioral clustering of school children. *Multivariate Behavioral Research*, 32, 105-134.
- \*Hulme, P. A., & Middleton, M. R. (2013). Psychosocial and developmental characteristics of civilly committed sex offenders. *Issues in Mental Health Nursing*, 34(3), 141-149.
- \*Hurt, S., & Oltmanns, T. F. (2002). Personality traits and pathology in older and younger incarcerated women. *Journal of Clinical Psychology*, 58(4), 457-464.
- Hutton, H. E., Miner, M. H., & Langfeldt, V. C. (1993). The utility of the Megargee-Bohn typology in a forensic psychiatric hospital. *Journal of Personality Assessment*, 60, 572-587.
- \*Iliceto, P., Pompili, M., Candilera, G., Rosafio, I., Erbuto, D., Battuello, M., ... Girardi, P. (2012). Temperament, insecure attachment, impulsivity, and sexuality in women in jail. *Journal of Forensic Nursing*, 8(1), 23-29.
- Joseph, N. & Benefield, N. (2012). A joint offender personality disorder pathway strategy: An outline summary. *Criminal Behaviour and Mental Health*, 22, 210-217.
- \*Kalichman, S. C. (1990). Affective and personality characteristics of MMPI profile subgroups of incarcerated rapists. *Archives of Sexual Behavior*, 19(5), 443-459.

- Kelsey, K. R., Rogers, R., & Robinson, E. V. (2015). Self-report measures of psychopathy: What is their role in forensic assessments? *Journal of Psychopathology and Behavioral Assessment, 37*, 380-391.
- \*Knust, S., & Stewart, A. L. (2002). Risk-taking behaviour and criminal offending: an investigation of sensation seeking and the Eysenck personality questionnaire. *International Journal of Offender Therapy and Comparative Criminology, 46*(5), 586-602.
- Koenraadt, F., Mooij, A., & Van Mulbregt, J. (2007). *The mental condition in criminal law: Forensic mental health assessment in a residential setting*. Amsterdam: Dutch University Press.
- Kool, M., & Aalders, H. (2014). Behandeling van persoonlijkheidsstoornissen in een forensische setting [Treatment of personality disorder in a forensic setting]. *Tijdschrift voor Psychotherapie, 40*(5), 360-375.
- \*Kroner, D. G., & Reddon, J. R. (1992). The Anger Expression Scale and State-Trait Anger Scale. Stability, reliability, and factor structure in an inmate sample. *Criminal Justice and Behavior, 19*(4), 397-408.
- Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine, 42*(9), 1879-1890.
- Lally, S. J. (2003). What tests are acceptable for use in forensic evaluations? A survey of experts. *Professional Psychology: Research and Practice, 34*(5), 491-498.
- \*Laulik, S., Allam, J., & Sheridan, L. (2007). An investigation into maladaptive personality functioning in Internet sex offenders. *Psychology, Crime & Law, 13*(5), 523-535.
- Leshem, R., & Glicksohn, J. (2007). The construct of impulsivity revisited. *Personality and Individual Differences, 43*, 681-691.
- Lilienfeld, S. O. (1999). The relation of the MMPI-2 Pd Harris-Lingoes subscales to psychopathy, psychopathy facets, and antisocial behavior: implications for clinical practice. *Journal of Clinical Psychology, 55*, 241-255.
- Lilienfeld, S. O., & Widows, M. (2005). *Professional manual for the Psychopathic Personality Inventory-Revised (PPI-R)*. Lutz, FL: Psychological Assessment Resources.
- Livesley, W. J., & Jackson, D. N. (2006). *Manual for the Dimensional Assessment of Personality Pathology*. Port Huron, MI: Sigma Press.
- Looper, K. J., & Paris, J. (2000). What dimensions underlie Cluster B personality disorders? *Comprehensive Psychiatry, 41*, 432-437.
- \*Magaletta, P. R., Faust, E., Bickart, W., & McLearn, A. M. (2014). Exploring clinical and personality characteristics of adult male internet-only child pornography offenders. *International Journal of Offender Therapy and Comparative Criminology, 58*(2), 137-153.
- \*Magyar, M. S., Edens, J. F., Lilienfeld, S. O., Douglas, K. S., Poythress, N. G., & Skeem, J. L. (2012). Using the Personality Assessment Inventory to predict male offenders' conduct during and progression through substance abuse treatment. *Psychological Assessment, 24*(1), 216-225.
- Martin, R., Watson, D., & Wan, C. K. (2000). A three-factor model of trait anger: Dimensions of affect, behavior, and cognition. *Journal of Personality, 68*(5), 869-897.
- \*McCreary, C., & Padilla, E. (1977). MMPI differences among black, Mexican-American and white male offenders. *Journal of Clinical Psychology, 33*(1), 171-177.
- McEwan, T. E., Davis, M. R., MacKenzie, R., & Mullen, P. E. (2009). Brief report: The effects of social desirability response bias on STAXI-2 profiles in a clinical forensic sample. *British Journal of Clinical Psychology, 48*, 431-436.

- McGrath, R. E., Mitchell, M., Kim, B. H., & Hough, L. (2010). Evidence for response bias as a source of error variance in applied assessment. *Psychological Bulletin*, *136*(3), 450-470.
- McGrath, R. E., Rashid, T., Hayman, J., & Pogge, D. L. (2002). A comparison of MMPI-2 high-point coding strategies. *Journal of Personality Assessment*, *79*, 243-256.
- Megargee, E. I. (1997). Using the Megargee MMPI-based classification system with the MMPI-2s of female prison inmates. *Psychological Assessment*, *9*, 75-82.
- Megargee, E. I., Carbonell, J.L., Bohn, M.J., Jr, & Sliger, G.L. (2001). *Classifying criminal offenders with the MMPI-2: The Megargee System*. Minneapolis, MN: University of Minnesota Press.
- Mei-Tal, M. (2002). The criminal responsibility of psychopathic offenders. *Israel Law Review*, *36*(2), 103-121.
- Mehrabian, A. & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality*, *40*, 525-543.
- Mevis, P. A. M. (2011). Opnieuw de TBS: Laat svp tijd en ruimte voor rustige en bestendige ontwikkeling in de goede richting [Once again TBS: Please leave time and space for calm and stable development in the right direction]. *Delikt en Delinkwnt*, *4*(26), 339-364.
- Miller, C. S, Kimonis, E. R., Otto, R. K., Kline, S. M., & Wasserman, A. L. (2012). Reliability of risk assessment measures used in sexually violent predator proceedings. *Psychological Assessment*, *24*(4), 944-953.
- Miller, J. D., & Lynam, D. (2001). Structural models of personality and their relation to antisocial behavior: A meta-analytic review. *Criminology*, *39*, 765-798.
- Miller, J. D. & Lynam, D. R. (2003). Psychopathy and the Five-Factor Model of Personality: A Replication and Extension. *Journal of Personality Assessment*, *81*(2), 168-178.
- Miller, J. D. & Lynam, D. R. (2015). Understanding Psychopathy Using the Basic Elements of Personality. *Social and Personality Psychology Compass*, *9*(5), 223-237.
- Miller, J. D., Lynam, D. R., Widiger, T. A., & Leukefeld, C. (2001). Personality disorders as extreme variants of common personality dimensions. Can the Five-Factor Model of personality adequately represent psychopathy? *Journal of Personality*, *69*, 253-276.
- Mills, J. F., Loza, W., & Kroner, D. G. (2003). Predictive validity despite social desirability: Evidence for the robustness of self-report among offenders. *Criminal Behaviour and Mental Health*, *13*, 140-150.
- Milton, J., McCartney, M., Duggan, C., Evans, C., Collins, M., McCarthy, L., & Larkin, E. (2005). Beauty in the eye of the beholder? How high security hospital psychopathically-disordered patients rate their own interpersonal behaviour. *Journal of Forensic Psychiatry & Psychology*, *16*, 552-565.
- Mooij, A. (2004). *Toerekeningsvatbaarheid. Over handelingsvrijheid* [Criminal responsibility. About freedom to act]. Amsterdam: Uitgeverij Boom.
- Morey, L. C. (1991). *Personality Assessment Inventory - Professional manual*. Lutz, FL: Psychological Assessment Resources, Inc.
- Morey, L. C., Hopwood, C. J., Markowitz, J. C., Gunderson, J. G., Grilo, C. M., McGlashan, T. H., ... Skodol, A. E. (2012). Comparison of alternative models for personality disorders, II: 6-, 8- and 10-year follow-up. *Psychological Medicine*, *42*, 1705-1713.
- Mulder, J. L., Dekker, R., & Dekker, P. H. (2004). *Kaufman intelligentietest voor adolescenten en volwassenen: Nederlandstalige bewerking van de Kaufman Adolescent and Adult Intelligence Test: Handleiding* [Kaufman intelligence test for adolescents and adults: Dutch adaptation of the Kaufman Adolescent and Adult Intelligence Test: Manual]. Leiden: PITS.

- Mulder, R. T., Newton-Howes, G., Crawford, M. J., & Tyrer, P. J. (2011). The central domains of personality pathology in psychiatric setting. *Journal of Personality Disorders*, 25(3), 364–377.
- Murrie, D. C., Boccaccini, M., Johnson, J., & Janke, C. (2008). Does interrater (dis)agreement on Psychopathy Checklist scores in Sexually Violent Predator trials suggest partisan allegiance in forensic evaluation? *Law and Human Behavior*, 32, 352–362.
- Nagtegaal, M. H., Van der Horst, R. P., & Schönberger, H. J. M. (2011). *Inzicht in de verblijfsduur van tbs-gestelden* [Insight into length of stay of TBS patients]. The Hague: WODC.
- Nentjes, L., Bernstein, D. P., Arntz, A., Slaats, M. E., & Hannemann, T. (2015). Theory of mind, social desirability, and unlikely symptom reporting in offenders with and without psychopathy. *The Journal of Nervous and Mental Disease*, 203(8), 596–603.
- Neumann, C. S., & Hare, R. D. (2008). Psychopathic traits in a large community sample: Links to violence, alcohol use, and intelligence. *Journal of Consulting and Clinical Psychology*, 76, 893–899.
- \*Newberry, M., & Shuker, R. (2012). Personality assessment inventory (PAI) profiles of offenders and their relationship to institutional misconduct and risk of reconviction. *Journal of Personality Assessment*, 94(6), 586–592.
- Nieberding, R., Gacono, C. B., Pirie, M., Bannatyne, L. A., Viglione, D. J., Cooper, B., ... Frackowiak, M. (2003). MMPI-2 based classification of forensic psychiatric outpatients: An explanatory cluster analytic study. *Journal of Clinical Psychology*, 59, 907–920.
- Niesten, I. J. M., Nentjes, L., Merckelbach, H., & Bernstein, D. P. (2015). Antisocial features and “faking bad”: A critical note. *International Journal of Law and Psychiatry*, 41, 34–42.
- Nijman, H., Cima, M., & Merckelbach, H. L. G. J. (2003). Nature and antecedents of psychotic patient’s crimes. *Journal of Forensic Psychiatry & Psychology*, 14, 542–553.
- Nijman, H., De Kruyk, C., & Van Nieuwenhuizen, C. (2002). Gedragsveranderingen tijdens tbs-behandeling [Behavioral change during enforced treatment]. *Tijdschrift voor Psychiatrie*, 44(5), 339–344.
- Nijman, H., De Kruyk, C., & Van Nieuwenhuizen, C. (2004). Behavioral changes during forensic psychiatric (TBS) treatment in the Netherlands. *International Journal of Law and Psychiatry*, 27(1), 79–85.
- NOMS (2011). *Working with personality disordered offenders: A practitioners guide*. Ministry of Justice: National Offender Management Service.
- Norlander, B., & Eckhardt, C. (2005). Anger, hostility, and male perpetrators of intimate partner violence: A meta-analytic review. *Clinical Psychology Review*, 25, 119–152.
- Osberg, T. M., & Poland, D. L. (2001). Validity of the MMPI-2 basic and Harris–Lingoes subscales in a forensic sample. *Journal of Clinical Psychology*, 57(12), 1369–1380.
- Overall, J. E. & Gomez-Mont, F. (1974). The MMPI-168 for psychiatric screening. *Educational and Psychological Measurement*, 34(2), 315–319.
- Pagan, J. L., Oltmanns, T. F., Whitmore, M. J., & Turkheimer, E. (2005). Personality disorder not otherwise specified: Searching for an empirically based diagnostic threshold. *Journal of Personality Disorders*, 19(6), 674–689.
- Palermo, G. B. (2007). New vistas on personality disorders and criminal responsibility. *International Journal of Offender Therapy and Comparative Criminology*, 51, 127–129.
- \*Panton, J. H. (1976). Personality characteristics of death-row prison inmates. *Journal of Clinical Psychology*, 32(2), 306–309.
- Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology*, 51(6), 768–74.

- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In: H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in a psychological and educational measurement* (pp. 49-46). Mahwah NJ: Erlbaum.
- \*Percosky, A. B., Boccaccini, M. T., Bitting, B. S., & Hamilton, P. M. (2013). Personality Assessment Inventory scores as predictors of treatment compliance and misconduct among sex offenders participating in community-based treatment. *Journal of Forensic Psychology Practice*, 13(3), 192-203.
- Pfohl, B., Blum, N., & Zimmerman, M. (2006). *Structured Interview for DSM-IV Personality (SIDP-IV)*. Arlington, VA: American Psychiatric Publishing, Inc.
- Ray, J. V., Hall, J., Poythress, N. G., Rivera-Hudson, N., & Lilienfeld, S. O. (2013). The relation between self-reported psychopathic traits and distorted response styles: A meta-analytic review. *Personality Disorders: Theory, Research, and Treatment*, 4(1), 1-4.
- \*Reith, G., Crockett, D., & Craig, K. (1975). Personality characteristics in heroin addicts and nonaddicted prisoners using the Edwards Personality Preference Schedule. *The International Journal of the Addictions*, 10(1), 97-112.
- Rice, M. E., & Harris, G. T. (1990). The predictors of insanity acquittal. *International Journal of Law and Psychiatry*, 13, 217-224.
- \*Roman, D. D., & Gerbing, D. W. (1989). The mentally disordered criminal offender: A description based on demographic, clinical and MMPI data. *Journal of Clinical Psychology*, 45(6), 983-990.
- \*Roy, A., Carli, V., Sarchiapone, M., & Branchey, M. (2014). Comparisons of prisoners who make or do not make suicide attempts and further who make one or multiple attempts. *Archives of Suicide Research*, 18(1), 28-38.
- \*Ruiz, M. A., Cox, J., Magyar, M. S., & Edens, J. F. (2014). Predictive validity of the personality assessment inventory (PAI) for identifying criminal reoffending following completion of an in-jail addiction treatment program. *Psychological Assessment*, 26(2), 673-678.
- Salekin, R. T., Worley, C. A., & Grimes, R. D. (2010). Treatment of Psychopathy: A Review and Brief Introduction to the Mental Model Approach for Psychopathy. *Behavioral Sciences and the Law*, 28, 235-266.
- \*Scott, N. A., & Conn, M. G. (1979). Correspondence of the MMPI and MMPI-168 among incarcerated female felons. *Journal of Personality Assessment*, 43(5), 473-478.
- Shorey, R. C., Brasfield, H., Febres, J., & Stuart, G. L. (2011). The association between impulsivity, trait anger, and the perpetration of intimate partner violence among women arrested for domestic violence. *Journal of Interpersonal Violence*, 26(13), 2681-2697.
- \*Shechory, M., Weiss, J. M., & Weinstain, R. (2013). Differentiating offenders by index offense and personality inventories: The characteristics of adult probationers in Israel. *International Journal of Offender Therapy and Comparative Criminology*, 57(3), 312-331.
- Spaans, M., Barendregt, M., Muller, E., Van der Meer, M., & Rinne, T. (2014). Gestandaardiseerde psychodiagnostiek in het Pieter Baan Centrum [Standardized psychodiagnostic assessment in the Pieter Baan Center]. *De Psycholoog*, juli-augustus, 32-41.
- Sparr, L. F. (2009). Personality disorders and criminal law: An international perspective. *Journal of the American Academy of Psychiatry and the Law*, 37, 168-181.
- Spielberger, C. D. (1988). *Manual for the State-Trait Anger Expression Inventory (STAXI)*. Odessa, FL: Psychological Assessment Resources, Inc.
- Spielberger, C. D., Jacobs, G., Russel, S., & Crane, R. S. (1983). Assessment of anger: The State-Trait Anger Scale. In: J. N. Butcher and C. D. Spielberger (Eds.), *Advances in personality assessment*, Vol. 2 (pp. 161-189). Hillsdale NJ: Lawrence Erlbaum Associates.

- Spielberger, C. D., Johnson, E. H., Jacobs, G. A., Krasner, S. S., Oesterle, S. E., & Worden, T. J. (1986). *The Anger Expression (AX) Scale*. Tampa: University of South Florida Human Resources Institute.
- StataCorp (2013). *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.
- Steinberg, L., Albert, D., Cauffman, E., Banich, M., Graham, S., & Woolard, J. (2008). Age differences in sensation seeking and impulsivity as indexed by behavior and self-report: Evidence for a dual systems model. *Developmental Psychology*, 44(6), 1764-1778.
- Sturup, J., Edens, J. F., Sörman, K., Karlberg, D., Fredriksson, B., & Kristiansson, M. (2013). Field reliability of the psychopathy checklist-revised among life sentenced prisoners in Sweden. *Law and Human Behavior*.
- Taft, C. T., O'Farrell, T., Torres, S. E., Panuzio, J., Monson, C. M., Murphy, M., & Murphy, C. M. (2006). Examining the correlates of psychological aggression among a community sample of couples. *Journal of Family Psychology*, 20, 581-588.
- Ter Horst, P., Jessen, A., Bogaerts, S., & Spreen, M. (2015). Behandelduur tot eerste toekenning tbs-verlof in een Nederlandse forensische psychiatrische kliniek: Dossieronderzoek naar invloed van patiëntkenmerken en delict [Clinical decision-making with regard to the granting of escorted leave for forensic patients detained by court order in a Dutch psychiatric clinic: Role of gender, disorder and the type of offence in the procedure]. *Tijdschrift voor Psychiatrie*, 57(5), 314-321.
- Timmerman, I. G., & Emmelkamp, P. M. (2001). The relationship between traumatic experiences, dissociation, and borderline personality pathology among male forensic patients and prisoners. *Journal of Personality Disorders*, 15(2), 136-149.
- \*Twomey, J. F., & Hendry, C. H. (1969). MMPI characteristics of difficult-to-manage federal penitentiary offenders. *Psychological Reports*, 24(2), 546.
- Ullrich, S., Borkenau, P., & Marneros, A. (2001). Personality disorders in offenders: Categorical Versus dimensional approaches. *Journal of Personality Disorders*, 15(5), 442-449.
- Uterwijk, J. (2000). *WAIS-III Nederlandstalige bewerking: Technische handleiding [WAIS-III Dutch adaptation: Technical manual]*. Lisse: Swets & Zeitlinger.
- \*Valliant, P. M., De Wit M., & Bowes, R. (2004). Cognitive and personality factors associated with assaultive and domestic offenders. *Psychological Reports*, 94(3), 1180-1184.
- \*Valliant, P. M., Gauthier, T., Pottier, D., & Kosmyna, R. (2000). Moral reasoning, interpersonal skills, and cognition of rapists, child molesters, and incest offenders. *Psychological Reports*, 86(1), 67-75.
- Van den Berg, A., & Oei, T. I. (2009). Een sociale behandeling van forensische patiënten [Social treatment of forensic patients]. In Oei, T. I. & Groenhuijsen, M. S. (Eds.), *Forensische psychiatrie en haar grensgebieden: Actualiteit, geschiedenis en toekomst* (pp. 357-383). Deventer: Kluwer.
- Van der Wolf, M. J. (2012). *TBS - veroordeeld tot vooroordeel: een visie na analyse van historische fundamenten van recente knelpunten, het systeem en buitenlandse alternatieven [TBS - sentenced to prejudice: A view and analysis of the historical foundations for recent bottlenecks, the system and foreign alternatives]*. Nijmegen: Wolf Legal Publishers.
- Van Gemmert, N., Van Schijndel, C., Gordeau, W., & Casanova, L. (2013). *Forensische zorg in getal 2008-2012 [Forensic care in numbers: 2008-2012]*. The Hague: Directie Bestuursondersteuning ASK, DJI.
- Van Kampen, D., & De Beurs, E. (2009). *DAPP-BQ: Dimensionale assessment van persoonlijkheids-pathologie inclusief screeningsversie [DAPP-BQ: Dimensional assessment of personality pathology including screening version]*. Amsterdam: Hogrefe.

- Van Kampen, D., De Beurs, E., & Andrea, H. (2008). A short form of the Dimensional Assessment of Personality Pathology-Basic Questionnaire (DAPP-BQ): The DAPP-SF. *Psychiatry Research*, 160(1), 115-128.
- Van Kordelaar, W. F. (2003). Delictclusters en hun signaalwaarde voor gedragsdeskundig onderzoek [Offense clusters and their signaling value for psychological assessment]. In T.I. Oei & M.S. Groenhuijsen (Eds.), *Actuele ontwikkelingen in de forensische psychiatrie* [Current developments in forensic psychiatry] (pp. 69-84). Deventer: Kluwer.
- Van Kuijk, Y. A. J. M. (2011). De TBS nog veiliger en toch korter? [TBS treatment more safe and still shorter?] *Sancties*, 2, 83-90.
- Verheul R. & Widiger, T. A. (2004). A meta-analysis of the prevalence and usage of the personality disorder not otherwise specified (PDNOS) diagnosis. *Journal of Personality Disorders*, 18, 309-319.
- Verwaaijen, S., & Polak, M. (2010). *Forensische zorg in perspectief: Een inhoudelijke toekomstvisie op de forensische zorg* [Forensic care in perspective: A substantive vision of the future of forensic care]. Amersfoort: LBHIV/ODO.
- Waij, E. M, Spronk, B. E. L., & Canton, W. J. (2010). De verlofpraktijk in de tbs [Leave practice in TBS]. *Sancties*, 5, 225-232.
- Walters, G. D. (1988). Assessing dissimulation and denial on the MMPI in a sample of maximum security, male inmates. *Journal of Personality Assessment*, 52(3), 465-474.
- Walters, G. D. (2006). Risk-appraisal versus self-report in the prediction of criminal justice outcomes: A meta-analysis. *Criminal Justice and Behavior*, 33 1-24.
- Walters, G. D. (2007). Predicting institutional adjustment with the Lifestyle Criminality Screening Form and the Antisocial Features and Aggression Scales of the PAI. *Journal of Personality Assessment*, 88(1), 99-105.
- \*Walls, R., McGlynn, F. D., & Tingstrom, D. H. (1977). An evaluation of three short forms extracted from the group form MMPI responses in incarcerated offenders. *Journal of Clinical Psychology*, 33(2), 431-435.
- Ward, J. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58, 236-244.
- Warren, F., McGauley, G., Norton, K., Dolan, B., Preedy-Fayers, K., Pickering, A., & Geddes, J. R. (2003). *Review of treatments for severe personality disorder. Home Office Online Report 30/03*. Home Office: Research Development and Statistics Directorate.
- Warren, J. I., Murrie, D. C., Chauhan, P., Dietz, P. E., & Morris, J. (2004). Opinion formation in evaluating sanity at the time of the offense: an examination of 5175 pre-trial evaluations. *Behavioral Sciences & the Law*, 22, 171-186.
- Weisberg, Y. J., DeYoung, C. G., & Hirsh, J. B. (2011). Gender differences in personality across the ten aspects of the Big Five. *Frontiers in Psychology*, 2, 178.
- Wells, G. A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (2015). *The Newcastle-Ottawa Scale (NOS) for assessing the quality non-randomised studies in meta-analyses*. Ottawa Healthcare Institute.
- Whyte, S., Fox, S., & Coxell, A. (2006). Reporting of personality disorder symptoms in a forensic inpatient sample: Effects of mode of assessment and response style. *The Journal of Forensic Psychiatry & Psychology*, 17(3), 431-441.
- Widiger, T. A. (2000). Personality Disorders in the 21st Century. *Journal of Personality Disorders*, 14(1), 3-16.

- Widiger, T. A., & Costa, P. T. (2012). Integrating Normal and Abnormal Personality Structure: The Five-Factor Model. *Journal of Personality, 80*, 1471-1506.
- Widiger, T. A., & Samuel, D. B. (2005). Evidence-Based Assessment of Personality Disorders. *Psychological Assessment, 17*(3), 278-287.
- Wygant, D. B., & Lareau, C. R. (2015). Civil and criminal forensic psychological assessment: Similarities and unique challenges. *Psychological Injury and Law, 8*(1), 11-26.
- Zimmerman, M., & Coryell, W. H. (1990). Diagnosing personality disorders in the community: A comparison of self-report and interview measures. *Archives of General Psychiatry, 47*, 527-531.
- Zimmerman, M., Rothschild, L., & Chelminski, I. (2005). The prevalence of DSM-IV personality disorders in psychiatric outpatients. *American Journal of Psychiatry, 62*, 1911-1918.

## References from the current dissertation

- Spaans, M., Barendregt, M., Haan, B., Nijman, H., & De Beurs, E. (2011). Diagnosis of antisocial personality disorder and criminal responsibility. *International Journal of Law and Psychiatry*, 34, 374-378.
- Spaans, M., Barendregt, M., Muller, E., De Beurs, E., Nijman, H., & Rinne, T. (2009). MMPI profiles of males accused of severe crimes: a cluster analysis. *Psychology, Crime & Law*, 15, 441-450.
- Spaans, M., De Beurs, E., Rinne, T., & Spinhoven, P. (in press). Prognostic factors for length of enforced forensic treatment with special focus on personality disorder. *International Journal of Forensic Mental Health*.
- Spaans, M., Molendijk, M. L., De Beurs, E., Rinne, T., & Spinhoven, P. (2016). Self-reported personality traits in forensic populations: A meta-analysis. *Psychology, Crime & Law*. Advance online publication.
- Spaans, M., Rinne, T., De Beurs, E., & Spinhoven, P. (2015). The DAPP-SF as a screener for personality disorder in a forensic setting. *Journal of Personality Assessment*, 97, 172-181.

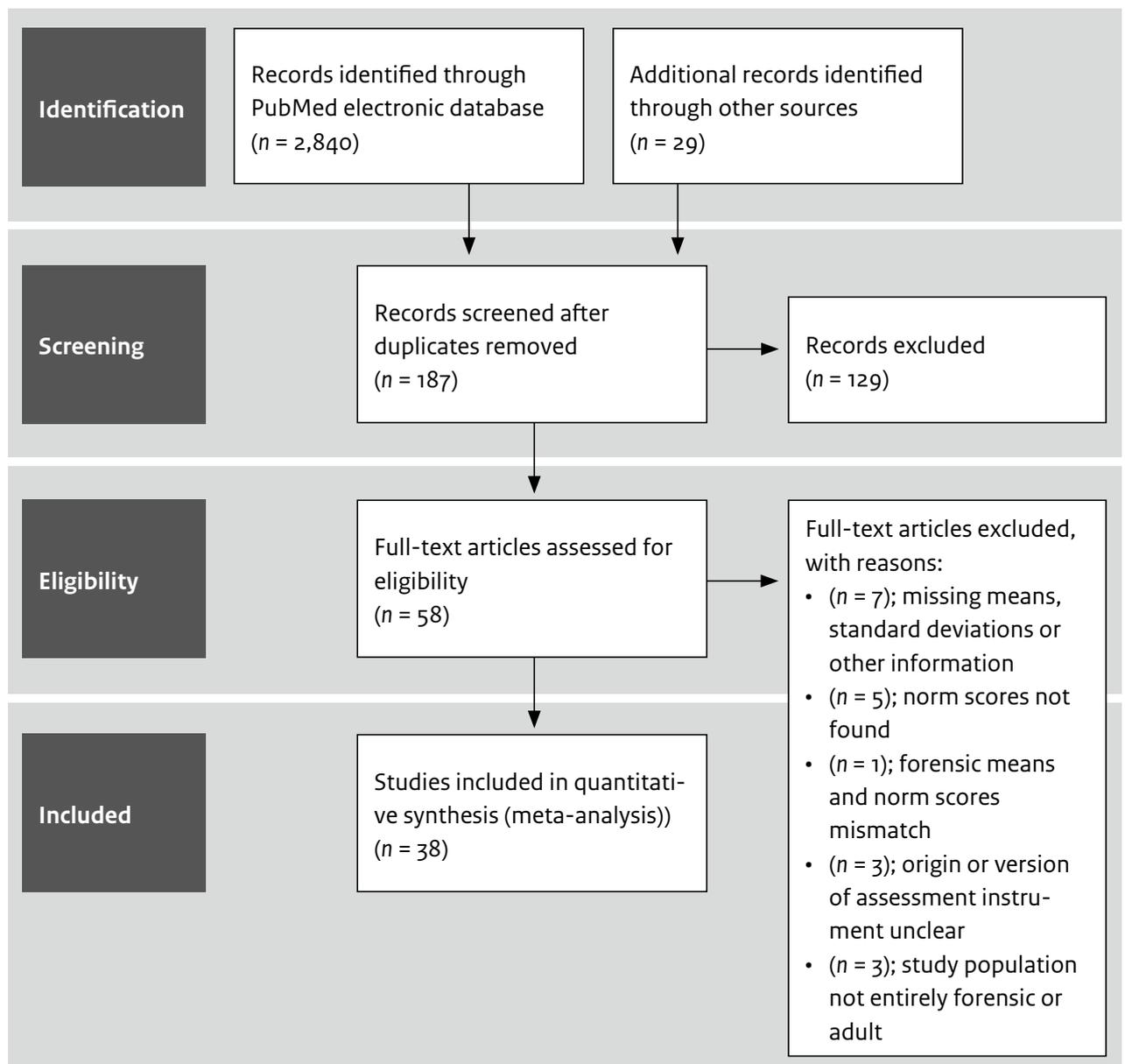
# Appendix 1

## Search terms for meta-analysis and systematic review

Prisons[MeSH] OR Prisoners[MeSH] OR Incarcerat\*[ti] OR Probati\*[ti] OR Prison\*[ti] OR Imprison\*[ti] OR Jail\*[ti] OR Inmat\*[ti] OR Penitent\*[ti] OR Detention\*[ti] OR Detain\*[ti] OR Probati\*[ti] OR Incarcerat\*[ti] OR Gaol\*[ti] OR ((Penal\*[ti] OR Correct\*[ti]) AND (Institut\*[ti] OR System\*[ti])) AND (Personality[MeSH] OR Mental Health[MeSH] OR Personality[ti] OR Personalities[ti] OR Assertiven\*[ti] OR Authoritarianis\*[ti] OR Character\*[ti] OR Creativity[ti] OR Dependency[ti] OR Empath\*[ti] OR Individuality[ti] OR Intelligence[ti] OR IQ[ti] OR Leadership\*[ti] OR Machiavellianis\*[ti] OR Negativis\*[ti] OR Ego[ti] OR Extravers\*[ti] OR Identification[ti] OR Identity[ti] OR Identities[ti] OR Individuation[ti] OR Introversi\*[ti] OR "Moral Development"[ti] OR "Psychosexual Develop\*" [ti] OR "Self Concept\*" [ti] OR Superego[ti] OR Unconscious[ti] OR Temperament\*[ti] OR Mental Health\*[ti]

## Appendix 2

### Flow diagram of search process for meta-analysis and systematic review



## Dutch summary

In de meeste landen kan iemand die een misdaad pleegt wegens een stoornis die de wilsvrijheid beïnvloedt, worden beschouwd als niet of minder strafbaar. De dader kan dan helemaal worden vrijgesteld van strafrechtelijke vervolging, strafvermindering zoals een lichtere gevangenisstraf krijgen of verplicht worden behandeld. In Nederland vallen persoonlijkheidsstoornissen hier ook onder en kan de aanwezigheid van zo'n stoornis ten tijde van het delict reden zijn voor verminderde toerekeningsvatbaarheid of behandeling onder dwangverpleging (Barendregt et al., 2008; De Kogel et al., 2006; Nijman et al., 2004). Het Nederlandse systeem van gedwongen behandeling van (persoonlijkheids-)gestoorde daders heeft de afgelopen jaren sterk onder vuur gelegen naar aanleiding van een aantal ernstige incidenten. Het aantal forensisch psychiatrische patiënten in Forensisch Psychiatrische Centra (FPC's) is de afgelopen jaren sterk toegenomen (van 405 patiënten 1990 tot meer dan 2,100 in 2010). Ook heeft het systeem een groot aantal beleidsveranderingen ondergaan, maar hebben deze veranderingen elkaar zo snel opgevolgd dat veel beleid al weer was teruggedraaid voordat het konden worden geëvalueerd (Nagtegaal et al., 2011). Volgens Nijman en collega's (2004), verblijven er nu meer psychiatrisch forensische patiënten met een persoonlijkheidsstoornis dan met psychotische stoornissen in FPC's. Daarom is het essentieel dat persoonlijkheidspathologie – bestaande uit zowel persoonlijkheidsstoornissen als problematische persoonlijkheidskenmerken – grondig en nauwkeurig vastgesteld wordt binnen de forensische gezondheidszorg.

## Gestandaardiseerde psychodiagnostiek in het Pieter Baan Centrum

Het Pieter Baan Centrum (PBC) is de psychiatrische observatiekliniek van het Nederlandse Ministerie van Veiligheid en Justitie waar jaarlijks zo'n 220 rapportages worden opgemaakt over verdachten van ernstige misdrijven. Een PBC-rapportage wordt opgesteld door een multidisciplinair team, bestaande uit o.a. een psychiater, psycholoog, een groepsleider op de verblijfsafdeling, een zogenaamde milieuonderzoeker die de levensloop van de verdachte uitzoekt en een jurist en geeft de rechter aan of er ten tijde van het delict een stoornis aanwezig was bij de verdachte, wat de mate van toerekeningsvatbaarheid was en of behandeling wel of niet onder dwangverpleging geïndiceerd is. Het advies van het PBC wordt in 86% van de gevallen overgenomen door de rechter (Boonekamp et al., 2008).

Testpsychologisch onderzoek is een belangrijke bron van informatie tijdens een PBC-onderzoek. In 2007 is het diagnostisch proces van het PBC kritisch bekeken en herzien om te komen tot meer uniformiteit en standaardisatie en een toename van het gebruik van wetenschappelijk onderbouwde en betrouwbare meetinstrumenten (Spaans et al., 2014). Een gunstig bijeffect van deze standaardisatie was de opbouw van een digitaal gegevensbestand van testcores van de forensisch onderzochte verdachten waarmee onderzoek gedaan kon worden naar de geselecteerde meetinstrumenten. Dit was van belang omdat de meeste instrumenten die in de forensische gezondheidszorg beschikbaar en gebruikt zijn niet specifiek voor die populatie zijn ontwikkeld (Wygant & Lareau, 2015). Met het opgebouwde gegevensbestand kon de bruikbaarheid van deze instrumenten binnen de forensische

gezondheidszorg nader bestudeerd worden. De resultaten van dat onderzoek worden weergegeven in dit proefschrift, dat de prevalentie, diagnostiek en prognostische waarde voor de behandeling van persoonlijkheidspathologie binnen de forensische gezondheidszorg onderzoekt.

## Prevalentie van persoonlijkheidspathologie in de forensische gezondheidszorg

Persoonlijkheidspathologie is zeer prevalent in gevangenispopulaties: zo'n 65% van de mannen en 42% van de vrouwen in westerse gevangenissen heeft een persoonlijkheidsstoornis (Fazel & Danesh, 2002). Persoonlijkheidsstoornissen uit cluster B van de Diagnostic Statistical Manual (DSM) van de American Psychiatric Association (APA) komen hierbij het meest voor: antisociaal, borderline, teatraal en narcistisch (Fazel & Danesh, 2002; De Ruiter & Greeven, 2000; Hildebrand & De Ruiter, 2004; Timmerman & Emmelkamp, 2001) alsook de bijbehorende persoonlijkheidstrekken zoals een hoge mate van agressie of vijandigheid, slechte impulscontrole, het zoeken van sensatie en gebrek aan empathie (American Psychiatric Association, 2000; Boccaccini et al., 2010; Cunradi et al., 2009; Dolan & Blackburn, 2006; Hare, 2006; Looper & Paris, 2000; Norlander & Eckhardt, 2005; Shechory et al., 2011; Walters, 2007).

De diagnose van persoonlijkheidspathologie is in de forensisch gezondheidszorg complexer dan in de reguliere psychologische praktijk. Ten eerste hebben de meeste verdachten niet zelf om een evaluatie gevraagd en is er doorgaans sprake van zeer weinig vertrouwen tussen de psycholoog en de verdachte (Cima, 2013). Ten tweede hebben verdachten en patiënten in de forensische gezondheidszorg heel wat te winnen of verliezen op basis van de resultaten van hun beoordeling, zoals de duur van hun gevangenisstraf of een behandeling onder dwang. Kwesties als simulatie (het overdrijven van negatieve eigenschappen) of dissimulatie (het bagatelliseren van negatieve eigenschappen) kunnen daarom zeker een rol spelen in de manier waarop een individu zichzelf presenteert (Cima, 2013; Wygant & Lareau, 2015). De aanwezigheid van een (cluster B) persoonlijkheidsstoornis of psychopathische persoonskenmerken kunnen eveneens vertekende resultaten geven (Cima, 2003)

## Diagnostiek bij persoonlijkheidspathologie in de forensische gezondheidszorg

Persoonlijkheidspathologie kan op meerdere wijzen worden vastgesteld, zoals door middel van zelfrapportage, beoordelingschalen en (semi-)gestructureerde interviews. Hiervan is zelfrapportage het minst ingewikkeld en tijdrovend voor de forensische gedragsdeskundigen aangezien deze instrumenten snel en eenvoudig af te nemen zijn en de meeste inspanning vereisen van de beoordeelde, niet de beoordelaar.

Toch zijn er twijfels geuit over de validiteit van zelfrapportage binnen de forensische gezondheidszorg i.v.m. mogelijke afwijkende responsstijlen als (dis)simulatie (Milton et al., 2005). Gezien de grote diversiteit van zelfrapportage-instrumenten die beschikbaar zijn voor persoonlijkheidskenmerken – waarvan de meerderheid niet specifiek voor de forensische gezondheidszorg is ontworpen – lijkt er

weinig duidelijkheid en overzicht te bestaan van de huidige kennis over deze kenmerken in het forensische veld. Hoofdstuk Twee presenteert een onderzoek naar de prevalentie van zelfgerapporteerde forensisch relevante persoonlijkheidskenmerken bij gedetineerden wereldwijd. Er werd een systematische review en meta-analyse uitgevoerd van zelf-gerapporteerde woede, agressie, vijandigheid, antisociale trekken, psychopathie en impulsiviteit. De scores van gedetineerden werden vergeleken met die van niet-gedeteneerde norm- of referentiegroepen.

Er werden geen algemene verschillen gevonden in zelfgerapporteerde niveaus van woede, agressie, vijandigheid, of impulsiviteit. Zelfgerapporteerde niveaus van antisociale en psychopathische kenmerken waren in de forensische gezondheidszorg aanzienlijk en significant hoger dan in de referentiegroepen. Dit is in lijn met soortgelijke studies die vonden dat forensische patiënten met antisociale neigingen of psychopathische kenmerken naar waarheid antwoorden op vragen over deze eigenschappen (Cima, Van Bergen, & Kremer, 2008; Edens, 2009; Niesten et al., 2015; Ray et al., 2013). Tegelijkertijd suggereren andere studies dat zelfrapportage een onderschatting van negatieve persoonlijkheidskenmerken oplevert (Blackburn et al., 2004; De Ruiter & Greeven, 2000; Hornsveld et al., 2009; Milton et al., 2005; Spaans et al., 2015; Zimmerman & Coryell, 1990).

Een alternatieve verklaring voor de resultaten is de relatieve oververtegenwoordiging van verschillende versies van de MMPI bij de gevonden studies. De Psychopathic Deviate (Pd) schaal van de MMPI meet niet zozeer psychopathische of antisociale kenmerken maar problemen met de interpersoonlijke relaties, impulscontroleproblemen en vervreemding (Butcher et al., 2015). Een andere verklaring voor de resultaten is de keuze voor drie overkoepelende categorieën van forensisch relevante persoonlijkheidskenmerken op basis van de klinische ervaring van de auteurs en niet van eerder onderzoek of gevalideerde constructen.

Het blijft onduidelijk in welke mate en voor welke persoonlijkheidskenmerken kwesties als (dis) simulatie een rol spelen. Er kan ook sprake zijn van onbedoelde zelfdeceptie (Ray et al., 2013). In ieder geval geven de resultaten van Hoofdstuk Twee aan dat men voorzichtig dient te zijn bij het gebruik van zelfrapportage in het forensische veld, vooral wanneer deze instrumenten niet beschikken over een schaal die afwijkende responsstijlen zoals sociale wenselijkheid meet.

Om verschillende redenen is het belangrijk dat het forensische psychodiagnostische proces efficiënt en kosteneffectief is (Wygant & Lareau, 2015). Binnen de reguliere psychologische praktijk raden Widiger en Samuel (2005) een tweetrapsaanpak aan voor een efficiënte beoordeling van persoonlijkheidspathologie. Eerst wordt een zelfrapportagevragenlijst voorgelegd om te screenen op de mogelijke aanwezigheid van persoonlijkheidsstoornissen, waarna er in geval van indicaties van een stoornis een (semi-)gestructureerd interview afgenomen wordt om de aanwezigheid hiervan te toetsen. Een snelle en efficiënte screeningsvragenlijst kan veel (tijds)voordelen opleveren voor de forensische gezondheidszorg. Tegelijkertijd blijft de belangrijkste eis van zelfrapportage-instrumenten volgens Wygant en Lareau (2015) dat ze nauwkeurig zijn en dienen forensische gedragskundige rapporten duidelijk, objectief en transparant te zijn.

Hoofdstuk Drie bespreekt de geschiktheid van de Dimensional Assessment of Personality Pathology – Short Form (DAPP-SF; Van Kampen et al., 2008) voor deze tweetrapsaanpak in een psychiatrische observatiekliniek waar verdachten van ernstige misdrijven een psychologisch en psychiatrisch onderzoek ondergaan. De DAPP-SF is een zelfrapportage-instrument voor de persoonlijkheid die betrouwbaar en valide is gebleken in de algemene bevolking, bij patiënten die behandeld worden voor persoonlijkheidsstoornissen (Van Kampen et al., 2008) en bij patiënten met stemmings-, angst- en somatoforme stoornissen (De Beurs et al., 2009). Bovendien kan de DAPP-SF patiënten met

persoonlijkheidsstoornissen onderscheiden van de algemene bevolking (De Beurs et al., 2010). Aangezien dit instrument op die manier een succesvolle screener is binnen de algemene populatie, kan het dat wellicht ook zijn binnen de forensische gezondheidszorg.

De verwachting was dat wanneer een of meer (sub)schalen van de DAPP-SF hoge scores vertoonden, er een of meerdere persoonlijkheidsstoornissen zouden worden geïdentificeerd door middel van het Structured Interview for DSM-IV Personality (SIDP-IV). De resultaten toonden echter dat de forensische populatie beduidend lagere persoonlijkheidspathologie rapporteerde dan de algemene populatie op 14 van de 22 persoonlijkheidsschalen van de DAPP-SF. Op basis van dit bodemeffect en andere onopvallende resultaten werd geconcludeerd dat de bruikbaarheid van de DAPP-SF als screener voor persoonlijkheidsstoornissen binnen de context van de forensische gezondheidszorg zeer beperkt is.

Opnieuw rijst de vraag of de mate van zelfgerapporteerde persoonlijkheidspathologie overeenkomt met de werkelijke niveaus of dat de respondenten (dis)simuleerden tijdens het invullen van de vragenlijst om bijvoorbeeld hun kansen op (voor hen) ongewenste juridische gevolgen te verminderen, zoals gedwongen opname in een beveiligd forensisch psychiatrisch centrum (FPC).

Nogmaals terugkomend op het onderwerp van zelfrapportage-instrumenten in het forensische veld is ook de Minnesota Multiphasic Personality Inventory-2 (MMPI-2; Hathaway & McKinley, 1989; Nieberding et al., 2003) een interessant instrument. De MMPI-2 is een van de meest gebruikte zelfrapportage-instrumenten in de forensische gezondheidszorg, ook al is het instrument niet gevalideerd voor forensische populaties. Een populaire onderzoekslijn met de MMPI(-2) binnen de forensische gezondheidszorg is het indelen van daders in verschillende profielgroepen op basis van hun scores. Dergelijke classificaties kunnen nuttig zijn voor de behandeling van elk specifieke subtype daders. Eerdere studies vonden tussen de twee (Espelage et al., 2003; Hall et al., 1991) en tien (Megargee et al., 2001) verschillende clusters. Studies waaruit slechts twee aparte clusters blijken leiden tot twijfels over de geschiktheid van de MMPI-2 voor het differentiëren tussen persoonlijkheidsprofielen in het veld van de forensische gezondheidszorg. Daarom werd het clusteranalyse-onderzoek gerepliceerd bij verdachten van ernstige misdrijven die psychiatrische observatie ondergingen. Uit de resultaten, weergegeven in Hoofdstuk Vier, bleken slechts twee clusters van verschillende persoonlijkheidsprofielen: één die een afwezigheid van psychische problemen rapporteerde en één die verhogingen op alle behalve drie schalen rapporteerde.

Geconcludeerd werd dat verschillende profielen niet effectief werden onderscheiden door de MMPI-2. Opnieuw rijzen er twijfels over het gebruik van zelfrapportage in forensische populaties en vooral de bruikbaarheid van de MMPI-2 voor diagnostische doeleinden. De resultaten suggereren dat dat, ondanks de populariteit van het instrument, binnen een forensische context beperkt is tot het screenen van de aan- of afwezigheid van algemene psychopathologie en dat het instrument niet goed in staat is om onderscheid te maken tussen subtypen daarin.

## Persoonlijkheidsstoornis en het strafrecht

Wanneer persoonlijkheidspathologie eenmaal is vastgesteld bij verdachten van ernstige misdrijven, is het aan de rechter om een uitspraak te doen over de mate van toerekeningsvatbaarheid en een eventuele straf en/of behandeling. Vanwege verschillende bevindingen over de vraag of persoonlijkheidsstoornissen wel of niet door cognitieve of neurologische tekorten tot stand komen en op die manier de wilsvrijheid beïnvloeden (Ciocchetti, 2003; Fine & Kennett, 2004; Herpertz & Sass,

2000; Mei-Tal, 2002; Palermo, 2007; Sparr, 2009) en de complexe rol die persoonlijkheidsstoornis internationaal gezien al dan niet speelt bij het bepalen van toerekeningsvatbaarheid, bestudeert het onderzoek in Hoofdstuk Vijf de Nederlandse forensische context. Er werden twee studies uitgevoerd naar de vraag of persoonlijkheidsstoornissen en psychopathische trekken bij verdachten van ernstige misdrijven volgens Nederlandse forensische gedragsdeskundigen redenen zijn voor verminderde toerekeningsvatbaarheid en/of behandeling onder dwang. De resultaten toonden een internationaal unieke rol van persoonlijkheidsstoornis in Nederland, namelijk dat het leidt tot verminderde toerekeningsvatbaarheid en een advies voor behandeling onder dwang. De resultaten toonden ook dat van kenmerken die betrekking hadden op impulsiviteit en (on)verantwoordelijkheid wordt gedacht dat ze de wilsvrijheid beïnvloeden. De resultaten van deze studie tonen aan dat Nederlandse forensische deskundigen de aanwezigheid van een persoonlijkheidsstoornis of psychopathie zorgvuldig meewegen in hun adviezen over toerekeningsvatbaarheid en eventuele gedwongen opname in een FPC. Dit benadrukt het belang van een grondige en zorgvuldige diagnose van de persoonlijkheidspathologie bij psychologische en psychiatrische beoordelingen van verdachten in strafzaken.

## Persoonlijkheidsstoornis en behandeling onder dwangverpleging

In Nederland, wanneer de rechtbank heeft vastgesteld dat een psychische stoornis de wilsvrijheid beïnvloedde en leidde tot een misdaad, kan de toerekeningsvatbaarheid als verminderd beschouwd worden en kan de dader de maatregel van gedwongen opname in een FPC worden opgelegd. Het doel van de behandeling is om de samenleving te beschermen tegen personen met een hoog risico op recidive (De Kogel et al., 2006; Van Gemmert et al., 2013) en om een geleidelijke en veilige terugkeer in de maatschappij van het individu te begeleiden (Koenraadt et al., 2007). Zolang het recidiverisico van de forensisch psychiatrische patiënt niet voldoende is afgenomen om terugkeer in de maatschappij te rechtvaardigen, wordt de behandeling voortgezet waardoor levenslange opname op een long-stay afdeling mogelijk wordt. De mediaan behandelduur nam door de jaren heen geleidelijk toe, van 7,3 jaar voor het instroomcohort van 1990 (Dienst Justitiële Inrichtingen, 2011) tot een piek van 10,6 jaar voor het instroomcohort van 1998 (Dienst Justitiële Inrichtingen, 2015). Een recente studie toonde aan dat de mediaan behandelduur voor patiënten die in 2006 instroomden enigszins is afgenomen tot 8,0 jaar (Dienst Justitiële Inrichtingen, 2015).

Begeleid en onbegeleid verlof zijn cruciale stappen in het behandelingsproces. Deze periodes van geautoriseerde vrijheid om zich buiten de beveiligde zone van de kliniek te bewegen, bereiden de patiënt geleidelijk voor op zijn of haar terugkeer in de maatschappij, door te onderzoeken of de vooruitgang die in de kliniek is geboekt, standhoudt in de minder beschermende buitenomgeving en of de patiënt de toename van vrijheid en verantwoordelijkheid aankan (Nagtegaal et al., 2011; Bernstein et al., 2012).

Het uitbreiden van de kennis over de persoonlijkheidsstoornis als voorspellende factor voor behandelduur kan bijdragen tot een meer doeltreffende forensische behandeling. Hoofdstuk Zes presenteert een prospectief onderzoek naar de voorspellende waarde van persoonlijkheidsstoornis voor behandelduur ten tijde van het eerste verlofverzoek en aan het einde van de behandeling van 536 forensisch psychiatrische patiënten in FPC's in Nederland. De prevalentie van persoonlijkheids-

stoornis bij deze groep was 76,5%. Ook werden de effecten van indexdelict, As I-stoornis, middelen-misbruik en intellectueel functioneren onderzocht voor mogelijk versturende effecten op de relatie tussen persoonlijkheidsstoornis en behandelduur.

De resultaten toonden geen onafhankelijke voorspellende waarde van de persoonlijkheidsstoornis voor de behandelduur wanneer As-I stoornis, middelenmisbruik en intellectueel functioneren ook mee werden genomen in de analyses. Factoren die wel invloed hadden op behandelduur waren een indexdelict van (poging tot) geweld en poging tot moord/doodslag – die in verband stonden met een kortere behandelduur – en een indexdelict van (poging tot) een zedendelict, dat in verband stond met een langere behandelduur. Uit een post-hoc exploratieve analyse van een subgroep van patiënten met alleen een diagnose voor persoonlijkheidsstoornis bleek dat deze groep een kortere behandelduur had dan patiënten met een persoonlijkheidsstoornis en een comorbide As I-stoornis.

## Beperkingen

Een beperking van de studies in dit proefschrift is hun generaliseerbaarheid. Alle behalve de eerste studie werden uitgevoerd binnen de Nederlandse forensische en juridische context waarin, zoals weergegeven in Hoofdstuk Vijf, persoonlijkheidsstoornis een internationaal unieke rol speelt. Hierdoor zijn de resultaten van deze studies niet één op één te generaliseren naar andere landen. Ook bestonden enkele studiepogingen alleen uit in de kliniek opgenomen, voldoende Nederlands sprekende, niet-psychotische, mannelijke verdachten van ernstige misdrijven, die instemden met het onderzoek, wat leidt tot de mogelijkheid van selectiebias. De resultaten zijn niet één op één generaliseerbaar naar ambulante patiënten, veroordeelde delinquenten, vrouwen, of zij die verdacht van of veroordeeld zijn voor minder ernstige misdrijven dan de PBC-populatie.

Een beperking van de meta-analyse in Hoofdstuk Twee is dat de formulering van de drie overkoepelende categorieën van forensisch relevante persoonlijkheidskenmerken wellicht ook anders benaderd had kunnen worden, omdat sommige persoonlijkheidskenmerken tot meer dan één domein kunnen behoren. Een andere indeling had mogelijk tot andere resultaten geleid. Een beperking van het onderzoek naar de bruikbaarheid van de DAPP-SF als screener voor persoonlijkheidsstoornissen, gepresenteerd in Hoofdstuk Drie, is dat niet kon worden vastgesteld of het veronderstelde dissimuleren daadwerkelijk bij de onderzoekspopulatie voorkwam. Een alternatieve verklaring voor de bevindingen van Hoofdstuk Drie zou gebrek aan statistische power van de kleine steekproef kunnen zijn. Een verdere beperking is het ontbreken van gegevens over de interbeoordelaarsbetrouwbaarheid van forensische gedragsdeskundigen in het PBC op de SIDP-IV.

Het gebruik van digitale databases in Hoofdstuk Zes introduceerde enkele beperkingen. Het was niet mogelijk om kenmerken van de patiënt ten tijde van de daadwerkelijke behandeling op te nemen in het onderzoek omdat deze gegevens niet digitaal beschikbaar waren. In sommige gevallen waren meer dan tien (behandel)jaren verstreken sinds de diagnose die ten tijde van de psychiatrische beoordeling van de verdachte was gemaakt. Tien jaar geleden was de diagnostiek ten aanzien van persoonlijkheidsstoornissen minder accuraat dan tegenwoordig. Naast de mogelijkheid van menselijke fouten bij het handmatig invoeren van gegevens, is een onderzoek ook afhankelijk van de mate van detail van een digitale database. Zo was het niet mogelijk om eventuele comorbiditeit binnen persoonlijkheidsstoornissen of verschillen in de ernst van de (persoonlijkheid)stoornissen te onderzoeken.

## Implicaties van de resultaten voor de klinische praktijk

Forensische deskundigen dienen te allen tijde bedacht te zijn op de mogelijkheid van opzettelijke (dis) simulatie of onopzettelijk zelfbedrog bij zelfrapportage, vooral wanneer er (juridische) stimulansen zijn voor bepaalde uitkomsten. Conclusies van forensische gedragskundige beoordelingen moeten worden gebaseerd op een combinatie van evaluatiemethoden: (semi-)gestructureerde klinische interviews, uitgebreide collaterale informatie, cognitieve taken en hetero-anamnestiche gegevens. Zelfrapportage kan van zekere waarde zijn voor individueel gebruik of voor de opsporing van mogelijke persoonlijkheidspathologie maar dient altijd in combinatie met de op bovengenoemde andere manieren verzamelde informatie te worden gewogen. Gebruik van de DAPP-SF en de MMPI-2 voor specifieke diagnostische doeleinden bij forensisch psychiatrische populaties moet worden vermeden. De DAPP-SF is beperkt bruikbaar als screener voor persoonlijkheidsstoornissen en persoonlijkheidsprofielen worden onvoldoende onderscheiden door de MMPI-2.

Ook zou er over de hele wereld meer standaardisatie van instrumenten en terminologie moeten komen bij de beoordeling van persoonlijkheidspathologie, zodat studieresultaten met betrekking tot persoonlijkheidskenmerken en -stoornissen in de forensische veld gemakkelijker met elkaar kunnen worden vergeleken. Demeest effectieve behandelmethoden voor daders met persoonlijkheidsstoornissen blijken langdurig en intensief en mogelijk zou een combinatie van verschillende therapeutische benaderingen elkaar kunnen versterken. Bovendien moet er een toename komen van de digitale registratie van de patiënt, de behandeling, en de risicokenmerken gedurende de behandeling in de forensische gezondheidszorg en aan het einde daarvan, teneinde zinvol wetenschappelijk onderzoek, een betere (risico)diagnostiek, minder incidenten, minder recidive en een kortere behandelduur mogelijk te maken (Brand et al., 2009).

Een afname van de angst voor tbs zou een toename betekenen in het aantal verdachten van ernstige misdrijven die daarvoor in aanmerking zouden kunnen komen en die zouden meewerken aan het NIFP-onderzoek en daarmee ook een toename in data over persoonlijkheidsproblematiek.

## Suggesties voor verder onderzoek

Er is een dringende behoefte aan een testbatterij die speciaal is ontworpen voor de forensische behandeling en voor een beknopt screeningsinstrument voor persoonlijkheidsstoornissen dat met succes individuen die verdere beoordeling vergen kan identificeren. Het is van essentieel belang om meer standaardisatie in terminologie en instrumentarium aan te brengen op het gebied van persoonlijkheids-kenmerken, om forensisch gedragskundige resultaten wereldwijd goed te kunnen vergelijken. Een eerste stap zou zijn om alleen nog maar instrumenten te gebruiken die zijn gevalideerd of speciaal ontworpen zijn voor populaties in de forensische gezondheidszorg. Hiermee valt al een groot aantal instrumenten af. Ook is gedetailleerd onderzoek naar afwijkende antwoordtendensen, zoals sociale wenselijkheid, binnen verschillende juridische contexten geïndiceerd, zoals bij veroordeelde delinquenten of forensisch psychiatrische patiënten in plaats van verdachten, om de mogelijke invloed van (juridische) stimulansen op antwoordtendensen te onderzoeken. Dit zou hopelijk duidelijk maken of het de status van verdachte is die de beperkte bruikbaarheid van de DAPP-SF als screener en het onvermogen van de MMPI-2 om persoonlijkheidsprofielen vast te stellen heeft veroorzaakt, of dat het de setting van de forensische gezondheidszorg in het algemeen is.

Teneinde een meer valide meting van niveaus van forensisch relevante persoonlijkheidskenmerken in forensisch psychiatrische populaties te verkrijgen, dient gebruik te worden gemaakt van alternatieve meetmethoden voor zelfrapportage zoals (semi-)gestructureerde interviews, beoordelingschalen of cognitieve taken. Ook dient onderzocht te worden of verschillende persoonlijkheidskenmerken verschillende soorten afwijkende responsstijlen oproepen. De *evidence-base* voor de beste benadering blijft beperkt. Enkele auteurs beweren dat behandelmethoden voor daders met persoonlijkheidsstoornissen langdurig en intensief dienen te zijn en dat mogelijk verschillende therapeutische benaderingen dienen te worden gecombineerd. Een toename van (details van) digitale registratie van patiënt-, behandel- en risicokenmerken is van belang. Omdat in het verleden veel studies methodologische tekortkomingen hebben getoond (Warren et al., 2003), moeten inspanningen worden gedaan om deze tekortkomingen te vermijden. Hoewel het een grote uitdaging kan zijn binnen de context van de forensische gezondheidszorg, moet toekomstig onderzoek bij voorkeur bestaan uit gerandomiseerde studies met controlegroepen en gedetailleerde informatie geven over de studiepopulatie en selectie criteria, de wijze waarop diagnoses zijn gesteld (die gevalideerd zou moeten zijn voor forensisch psychiatrische populaties), over de behandeling die is toegediend (zodat deze later mogelijk gerepliceerd kan worden) en over de keuze voor de (gestandaardiseerde) uitkomstmaat en statistische analyses. Studies moeten gebruik maken van grote steekproeven en controlegroepen, eventuele uitval van respondenten bespreken en groepen met As I- en As II-problematiek waar mogelijk gescheiden houden. Het verrichten van nader zinvol onderzoek naar behandel-effectiviteit van welke behandeling voor welke patiënt in welk stadium zal hopelijk bijdragen aan betere (risico)diagnoses, minder incidenten, kortere behandelduur en minder recidive.

## Conclusie

Samenvattend kan worden gezegd dat, hoewel persoonlijkheidspathologie zeer prevalent is binnen de forensische gezondheidszorg volgens diverse studies, Hoofdstuk Twee, Drie en Vier van dit proefschrift hebben aangetoond dat forensische populaties moeilijk te beoordelen zijn door middel van zelfrapportage-vragenlijsten vanwege de mogelijkheid van misleidende en manipulatieve antwoordtendensen alsook onbewust zelfbedrog. Zelfrapportage kan van zekere waarde zijn voor individueel gebruik en de opsporing van mogelijke persoonlijkheidspathologie, maar voorzichtigheid is geboden bij de interpretatie van de resultaten. Een grondige en nauwkeurige diagnose van persoonlijkheidspathologie is van groot belang in het forensische veld. De gevolgen van een psychiatrische of psychologische beoordeling kunnen een grotere invloed hebben op het leven van de onderzochte dan in de reguliere gezondheidszorg.

Hoofdstuk Vijf en Zes hebben aangetoond dat binnen de context van de Nederlandse forensische gezondheidszorg persoonlijkheidsstoornis verminderde toerekeningsvatbaarheid en behandeling binnen een juridisch kader rechtvaardigt en dat een kleine subgroep van forensisch psychiatrische patiënten met alleen een persoonlijkheidsstoornis een kortere behandelduur hebben dan patiënten met een persoonlijkheidsstoornis en een comorbide As I-stoornis. Deze laatste studie heeft een drietal factoren geïsoleerd die behandelduur kunnen verkorten (een indexdelict van (poging tot) geweld en (poging tot) moord/doodslag) of verlengen (een indexdelict van (poging tot) een zedendelict) en daarmee een focus kunnen zijn voor beleidsmakers en forensische deskundigen in het behandelproces.



# Acknowledgements

Ik wil graag een aantal mensen bedanken, zonder wie dit proefschrift niet mogelijk was geweest. Ten eerste mijn promotores en co-promotor. Philip, ik wil jou bedanken dat jij dit reeds opgestarte proefschrift hebt aangenomen! Ik heb me vanaf het eerste moment zeer welkom gevoeld en heb erg veel gehad aan jouw kritische maar tegelijkertijd zeer verhelderende vragen. Ondanks dat je meerdere keren met mini-sabbatical ging (en met jaloersmakend mooie foto's als screensaver op je computer terugkwam) was je de afgelopen jaren voor mij altijd zeer goed bereikbaar.

Edwin en Thomas: wie had gedacht, toen jullie mij eind 2006 samen ontvingen voor een sollicitatiegesprek in het Pieter Baan Centrum, dat wij zoveel jaren later met z'n drieën hier zouden zijn? Edwin, bedankt dat je ondanks jouw vertrek bij het NIFP bij mijn proefschrift betrokken bent gebleven en dat je zo vaak tijd voor dit proefschrift vrij hebt gemaakt. Thomas, ook jou wil ik bedanken voor jouw nauwe betrokkenheid en het meedenken vanuit de praktijk door de jaren heen, en de ruimte en tijd die ik heb gekregen om mijn proefschrift af te ronden.

Ik wil al mijn co-auteurs bedanken voor hun bijdragen, met in het bijzonder Marko en Marc voor hun statistische/meta-analytische ondersteuning. Cyril: bedankt voor de DJI-data en voor het zo nu en dan fungeren als helpdesk. Ook wil ik de stagiaires bedanken die geholpen hebben bij het doorpluizen van metershoge stapels PBC-dossiers en wetenschappelijk artikelen of die contact hebben gezocht met uitgevers van meetinstrumenten van over de hele wereld. Tale, zonder jou was ik nu nog naar literatuur aan het zoeken. Ik ben je eeuwig dankbaar voor jouw grondige en snelle service.

DOP-ertjes: Dank voor alle gezelligheid, zowel binnen als buiten het PBC! Ik ga jullie missen, laten we de theetjes en borrels erin houden! PBC-ers: jullie hebben mij jarenlang stapels stoffige dossiers heen en weer zien sjouwen of regelmatig aan jullie bureau gehad met het zoveelste verzoek om data. Dank voor de prettige samenwerking! W&O-ers: na een tijdelijk bestaan als eenzame wetenschapper was het een verademing om lotgenoten te hebben en erg gezellig om bij de afdeling te horen. Ik wens jullie alle goeds! Maaike, jou wil ik in het bijzonder bedanken voor je betrokkenheid en de inhoudelijke en praktische steun die jij mij de laatste jaren hebt gegeven. Marie en Jos: mijn dank is groot voor jullie tijd en kritische blik.

Mijn ouders en broer wil ik bedanken voor hun liefde en steun. Vrienden: ik hoop dat nu eindelijk duidelijk is waar ik al die jaren mee bezig was! Ook wil ik mijn allerliefste man Marten bedanken. Zonder zijn onvoorwaardelijke en onmisbare steun, deels vanaf de andere kant van de wereld, had ik dit project niet af kunnen ronden. Ik heb veel gehad aan zijn advies en zakelijke kijk. Zo lang als wij elkaar al kennen, ben ik aan het promoveren. Het zal wellicht even wennen zijn, de nieuwe status quo! Voor mijn lieve paranimfen: dank dat jullie op deze spannende dag aan mijn zijde willen staan!

# Curriculum Vitae

Marleen Spaans was born in Bangkok, Thailand in 1981. After obtaining an International Baccalaureate (IB) diploma at the American Community School in Cobham, Great Britain as well as an A-level Dutch in 1999 she started studying Psychology at Leiden University. In 2002 she completed an Honours Research Project on the topic of intellectual and metacognitive skills. In 2003 she received her Master's Degree in Developmental and Educational Psychology and continued on to study Criminology at Leiden University. In 2006 she received her second Master's Degree in Forensic Criminology. From March, 2007 to August, 2016, she worked as a scientific researcher for the Department of Research and Development of the Netherlands Institute for Forensic Psychiatry and Psychology (NIFP), part of the Dutch Ministry of Security and Justice. This division later became the Department of Research and Education. In April of 2013 she was accepted as an external PhD candidate at the Graduate School of Social and Behavioural Sciences of Leiden University. Marleen currently lives in Dubai, United Arab Emirates.







Dit is een uitgave van:

NIFP/DJI

Postbus 13369 | 3500 IJ Utrecht

T 0800 646 39 51

E [nifp@dji.minjus.nl](mailto:nifp@dji.minjus.nl)

oktober 2016 | 96200

[www.nifpnet.nl](http://www.nifpnet.nl)