



Psychopathic traits and change on indicators of dynamic risk factors during inpatient forensic psychiatric treatment

Martin Hildebrand ^{*,1}, Corine de Ruiter ^{**2}

Van der Hoeven Kliniek, Utrecht, The Netherlands
Maastricht University, The Netherlands

ARTICLE INFO

Available online 10 May 2012

Keywords:

PCL-R
Psychopathy
Forensic psychiatric treatment
Dynamic risk factors
Forensic mental health

ABSTRACT

The main objective of the present study was to investigate the impact of treatment on forensic psychiatric inpatients, examining changes on 22 indicators of five dynamic risk factors for violence (i.e., egocentrism, hostility, impulsivity, lack of insight, and negative distrustful attitudes), and to relate these potential changes to level of psychopathy assessed with the Hare Psychopathy Checklist – Revised (PCL-R). Also, we studied the relationship between psychopathy and treatment compliance, as indicated by the attendance rate of therapeutic activities. Eighty-seven male patients (due to missing data on at least one measure, sample size varies from 58 to 87; 42 patients have complete datasets) were administered a standardized psychological assessment battery (self-report inventories, performance-based personality test, observer ratings) upon admission (T1) and after on average 20 months of treatment (T2). Upon admission, psychopathy (median split, PCL-R score ≥ 22) was significantly related to a higher score on five of the 22 indicators of dynamic risk. The analyses showed no significant differences between psychopathic and non-psychopathic patients on the indicators of dynamic risk factors during 20 months of inpatient forensic psychiatric treatment. However, psychopaths showed the expected pattern of treatment noncompliance, compared to non-psychopaths. The clinical and research implications of these findings are discussed.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Concerns about the effectiveness of treatment of forensic psychiatric patients focus particularly on patients with the diagnosis of psychopathic personality disorder. Psychopathic offenders present a major challenge to treatment providers because of the complex nature of the disorder itself (e.g., Blair, 2006; Patrick, 2007), its comorbidity with other types of mental disorder (e.g., Hart & Hare, 1989; Hildebrand & de Ruiter, 2004), and its association with other criminogenic risks and needs, such as substance abuse (Hillege, Das, & de Ruiter, 2010). Also, psychopathic offenders are more likely to recidivate than non-psychopathic offenders after incarceration or hospitalization (Leistico, Salekin, DeCoster, & Rogers, 2008).

Given the compelling evidence for its reliability and validity, the Hare Psychopathy Checklist – Revised (Hare, 1991, 2003), a 20-item clinical rating scale for the assessment of psychopathy in research, clinical, and forensic settings, has emerged as the standard for the

assessment of psychopathy. The PCL-R uses a semi-structured interview and collateral information to measure personality traits and behaviors related to a widely accepted conception of psychopathy (Cleckley, 1976; Hare, 2003). At least initially, PCL-R items were considered to be underpinned by two distinct but correlated factors: The “selfish, callous, and remorseless use of others” factor (F1) reflects the affective and interpersonal aspects of the disorder, whereas the behavioral items coalesce to form the “chronically, unstable and antisocial lifestyle; social deviance” factor (F2; Hare, 1991, p. 76). Cooke and Michie (2001), using confirmatory factor analysis and item-response theory analysis, developed a three-factor hierarchical model, in which the traditional F1 (Hare, 1991) was reconceptualized as two distinct factors named ‘arrogant and deceitful interpersonal style’ (PCL-R items 1, 2, 4, and 5) and ‘defective affective experience’ (items 6, 7, 8, and 16). The third factor was named ‘impulsive and irresponsible behavioral style’ (items 3, 9, 13, 14, and 15), excluding items reflecting antisocial behavior. Subsequent confirmatory factor analytic studies have found support for a four-factor model when nearly all of the PCL-R items are used (Hare, 2003; Vitacco, Rogers, Neumann, Harrison, & Vincent, 2005). Hare (2003) presented a two-factor, four-facet model for psychopathy, which retains the original two factors, but then divides each factor into more specific facets. F1 subsumes the Interpersonal and Affective facets, whereas F2 is separated into the Lifestyle and Antisocial facets. Criminal conduct, which is removed in Cooke and Michie’s (2001) three-factor model, is reincorporated in the Antisocial facet of the

* Correspondence to: M. Hildebrand, PhD, Hessenweg 123-B, 3731 JG, De Bilt, The Netherlands. Tel.: +31 30 2400465.

** Correspondence to: C. de Ruiter, PhD, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Tel.: +31 43 388 1905.

E-mail address: mhildebrand@casema.nl (M. Hildebrand).

¹ Martin Hildebrand is now a consultant forensic psychologist (private practice).

² Professor of Forensic Psychology, Department of Clinical Psychological Science, Maastricht University, Maastricht, The Netherlands.

two-factor, four-facet model (Hare, 2003). Thus, Hare (2003) confirmed the presence of the three-factors described by Cooke and Michie, and retained the remaining variables in the fourth facet described as *antisocial* behavior. The PCL-R total score can range from 0 to 40; scores of 30 or higher are considered the criterion for psychopathy, although many scholars have recognized the value of lower cut-off points of 25 or 26 in certain settings and populations, particularly in Europe (e.g., Cooke, 1995; Grann, Långström, Tengström, & Stålenheim, 1998; Hildebrand, de Ruiter, & de Vogel, 2004).

On the basis of previous research investigating the relationship between PCL-R psychopathy and response to treatment, four general conclusions can be drawn (for a review, see D'Silva, Duggan, & McCarthy, 2004). First, psychopathic offenders engage in more disruptive behavior during treatment when compared to non-psychopathic offenders (Hare, Clarke, Grann, & Thornton, 2000; Heilbrun et al., 1998; Hildebrand, de Ruiter, & Nijman, 2004; Rice, Harris, & Cormier, 1992). Rice et al. (1992), for example, reported that psychopathic offenders received more placements in seclusion for violent or disruptive behavior, more negative entries in the clinical records for disruptive or counter-therapeutic behavior, and more referrals to an institutional disciplinary sub-program than non-psychopathic offenders.

Second, psychopathic offenders appear to be less likely to remain in treatment compared to non-psychopathic offenders (e.g., Ogloff, Wong, & Greenwood, 1990; Seto & Barbaree, 1999). For example, Ogloff et al. (1990) examined 80 adult male offenders who volunteered to attend a corrections-based therapeutic community (TC) program. It was found that the length of stay in the program was significantly shorter for psychopathic offenders ($n=21$) than for other offenders ($n=59$), mainly because they were more likely to be discharged from the program due to misbehavior or lack of motivation. Hare et al. (2000) found that (PCL-R) psychopaths were less likely than nonpsychopaths to complete vocational and educational programs and were more likely than non-psychopathic offenders to be fired from a prison job.

Third, several studies found that higher PCL-R scores are associated with lower scores on global measures of change (e.g., Hobson, Shine, & Roberts, 2000; Hughes, Hogue, Hollin, & Champion, 1997; Ogloff et al., 1990). For example, Hughes et al. (1997) found that higher PCL-R scores were associated with lower scores on a global measure of clinical change, primarily rated from in-therapy progress information provided by treatment staff. Ogloff et al. (1990) reported that patients in the high-psychopathy group received lower ratings by treatment staff (blind to PCL-R scores) on improvement than non-psychopathic offenders.

Finally, the effectiveness of certain types of treatment in reducing (violent) re-offending among PCL-R psychopaths is questionable, and some treatments may even have harmful effects (e.g., Hobson et al., 2000; Rice et al., 1992). Rice et al. (1992), for example, conducted a retrospective evaluation of a TC in a maximum-security institution (Oak Ridge) for mentally disordered offenders, matching TC subjects with assessment-only subjects. Briefly, the TC program was based on one developed by Jones (1956, 1968). It was largely peer operated and involved intensive group therapy for up to 80 h per week. The goal was to create an environment that fostered empathy and responsibility for peers. Follow-up at a mean of 10.5 years after discharge showed that TC subjects with a PCL-R score of ≥ 25 showed higher rates of recidivism, particularly violent recidivism, than a comparable *no-treatment control* group, suggesting that this particular type of treatment may have had an adverse impact on psychopathic offenders. Although considered innovative at the time, the Oak Ridge TC treatment program would not be considered appropriate according to the current research literature on "what works" with offenders (Andrews & Bonta, 1994; Cooke & Philip, 2001). Results of recent PCL-R based studies suggest that psychopaths are as likely to benefit from treatment as non-psychopaths. Skeem (2008), for example, found in a prospective study ($N=381$ male offenders mandated to residential drug treatment) that psychopaths who received intensive treatment

where over three times less likely to be rearrested at one year follow-up than psychopaths who received less intensive treatment. Although psychopathy was associated with disruptive behavior and less perceived progress during treatment, the PCL-R scores did not moderate the effect of treatment dose on re-arrest rates (Skeem, 2008). Olver and Wong (2009), in a retrospective study of 156 sex offenders, found that sex offenders who showed positive therapeutic responses were less likely to recidivate in violent and sexual crimes, and that this relationship was not moderated by PCL-R scores.

It can be concluded from the research cited above that, compared to non-psychopaths, psychopathic offenders place a significant burden on any treatment setting. Psychopathy is associated with higher levels of institutional misbehavior, premature termination of treatment, and, in several studies, post-treatment re-offending. However, there seems to be consensus among scholars that this does not warrant the conclusion that the disorder is immutable (Blackburn, 2001; D'Silva et al., 2004; Hare, 1998; Hemphill & Hart, 2002; Salekin, 2002; Wong, 2000). Most treatment studies are hampered by serious methodological problems and provide little guidance concerning what is or is not effective (e.g., Hart & Hare, 1997; Hemphill & Hart, 2002; Wong, 2000).

Andrews, Zinger, et al. (1990; see also Andrews & Bonta, 2003; Andrews, Bonta, & Hoge, 1990) cogently argued that effective treatment to reduce recidivism requires the targeting of appropriate risk factors in offenders, including psychopathic offenders. Drawn from the risk/need/responsivity model (Andrews, Zinger, et al., 1990), the essence of the risk principle is that treatment is most effective when delivered proportionally to the level of risk of the patient. Thus, higher risk cases, such as psychopathic patients, should receive more intensive services (i.e., multifaceted intervention of longer duration), whereas lower risk cases should receive less intervention. Risk level is defined as the overall probability of criminal offending that is determined by both the number and severity of risk factors. The need principle refers to the type of treatment targets and suggests that interventions should be geared toward those factors that are most closely related to the risk of criminal offending (i.e., criminogenic needs). Examples of criminogenic need domains include problematic family and marital relationships, substance abuse, emotional instability and pro-criminal attitudes (e.g., Andrews & Bonta, 1994). The responsivity principle, finally, concerns the delivery of treatment programs in a style and mode that is consistent with the competency and learning style of the offender. The latter principle emphasizes the importance of patient characteristics and conditions that promote or impede positive change.

Hanson (1998; see also Andrews & Bonta, 2003) identified two general types of risk factors: static and dynamic. Static risk factors are those that are shown to be statistically related to recidivism and unable to change through intervention, and hence, can not be considered promising targets for treatment. Dynamic risk factors (criminogenic needs in Andrews and Bonta's terminology), on the other hand, are characteristics statistically related to recidivism that can (in principle) change (e.g., impulsivity, hostility, or negative attitudes), and when changed, are expected to result in a decrease in recidivism (Monahan & Appelbaum, 2000). Meta-analyses of the offender recidivism literature (Bonta, Law, & Hanson, 1998; Gendreau, Little, & Goggin, 1996) clearly revealed that dynamic 'need' variables correlate both with general and violent recidivism as well as or better than static factors (see also Dempster & Hart, 2002). This further stresses the importance of targeting these factors for violence-reducing strategies.

2. The present study

The main objective of the present study is to measure treatment progress in a sample of Dutch male offenders involuntarily admitted to a forensic psychiatric hospital, by using indicators of dynamic risk

factors as treatment outcome criterion. A dynamic risk factor, for the purpose of this study, is defined as a variable that relates to violence, may fluctuate with time and circumstances, and can be changed as a result of deliberate intervention (Webster, Douglas, Belfrage, & Link, 2000). The logic is as follows: dynamic risk factors relate to actual violence; these dynamic factors are capable of indexing change; therefore, systematic change as measured by these factors is expected to be associated with violence risk reduction.

We examined potential change in the following risk factors: egocentricity/narcissism, hostility, impulsivity, lack of insight, and negative distrustful attitudes. These variables, while not exhaustive, cover a broad range of dynamic risk factors generally considered relevant to violent reoffending (e.g., Andrews & Bonta, 2003; Bonta et al., 1998; Gendreau et al., 1996; Kay, Wolkenfeld, & Murrill, 1988; Menzies & Webster, 1995; Rice, Harris, Quinsey, & Cyr, 1990; Zamble & Quinsey, 1997), and, at least in theory, are apt to change. Some of these dynamic risk factors (i.e., impulsivity, lack of insight, negative attitudes) are also represented in the Historical Clinical Risk Management-20 (HCR-20; Webster, Douglas, Eaves, & Hart, 1997), a structured professional judgment risk assessment instrument that has drawn considerable international attention. Indeed, several studies using the HCR-20 have suggested that these dynamic factors are related to violence (e.g., Douglas, Ogloff, Nicholls, & Grant, 1999; Strand, Belfrage, Fransson, & Levander, 1999; for a review of HCR-20 validation research, see Otto & Douglas, 2010). Furthermore, McNiel, Eisner, and Binder (2003) found that hostile attitudes (aggressive attributional styles) were a significant contributor to violence risk in psychiatric inpatients. Baumeister, Smart, and Boden (1996) established a link between excessively high self-esteem (egocentrism) and aggression/violence (see also Baumeister, Bushman, & Campbell, 2000; Bushman & Baumeister, 1998).

We hypothesized that (1) upon admission to the hospital, offenders with high PCL-R scores would show more pathology on indicators of dynamic risk than patients with low scores on the PCL-R, and that (2) patients with high PCL-R scores would show more limited improvement after two years of inpatient treatment than patients with low PCL-R scores on dynamic outcome criteria. It should be noted that, in the present study, psychopathy is considered as a responsivity factor: an individual attribute that affects the achievement of treatment goals.

An additional objective of this study is to investigate the relationship between psychopathy and objective measures of treatment compliance, i.e., the extent to which the patient actually participates in the assigned treatment program. It was hypothesized that psychopathy is significantly related to a lower level of involvement on indicators of treatment compliance, including the ratio of the number of attended to planned (individual) psychotherapy sessions, attendance at work, educational activities, creative arts, and sports.

3. Method

3.1. Setting and treatment model

The study was conducted at the Van der Hoeven Kliniek, a forensic psychiatric hospital located in Utrecht, a medium-sized city (300,000 inhabitants) in The Netherlands. Patients have been sentenced by criminal court to involuntary commitment because of (severely) diminished responsibility for the offense they committed. In terms of legal status, most patients are sentenced to a 'maatregel van terbeschikkingstelling' (TBS-order), a judicial measure which can be translated as 'disposal to be treated on behalf of the state'. In theory, a TBS-order is of indefinite duration. Initially imposed for two years, it may be extended for one- or two-year periods as the court re-evaluates the patient to determine whether the risk of (violent) recidivism is still too high and treatment needs to be continued. TBS involves involuntary admission to a specialized maximum-security

forensic psychiatric hospital aimed at motivating the patient to participate voluntarily in the treatment offered by the hospital. The motivation for clinical practice is that it is legally permitted to place a patient on a residential ward with fellow patients and to structure his daily life in such a way that it is almost impossible for him to avoid contact with members of the hospital staff. Neither on ethical nor on legal grounds can there be an escape from the obligation to participate in a therapeutic milieu in order to facilitate social contacts aimed at motivating the patient for treatment. However, patients are free to refuse, for example, participating in specific therapeutic activities, such as individual or group psychotherapy. Because of the fact that the TBS-order can be extended by the court as long as the patient poses a risk, refusal of treatment generally implies a prolonged stay in the secure hospital.

The majority of the patients in the hospital (about 90%; see Hildebrand & de Ruiter, 2004) suffer from one or more DSM-IV(-TR) Axis II personality disorders (American Psychiatric Association, 1994, 2000), very often with a co-morbid substance use disorder. Schizophrenia or other psychotic disorders are present only in a minority (17%) of patients (Hildebrand & de Ruiter, 2004). With regard to type of mental disorders, Dutch forensic psychiatric samples are more similar to a prison sample in North America than to a North American forensic psychiatric sample where most patients are diagnosed with schizophrenia or other psychotic disorders (de Ruiter & Hildebrand, 2003).

The Van der Hoeven hospital as a whole is organized as a therapeutic community. The general treatment aim is a reduction in future violence risk by means of a positive change in those risk factors that are associated with (sexual) violence, and/or in protective factors that are expected to buffer the effect of risk factors for the individual patient. A central concept in the treatment model of the hospital is the stimulation of the patient's awareness that he is responsible for his own life, including his offenses and his progress in treatment. Patients reside in living groups of around 10 patients, where they can develop and practice new interpersonal styles and skills. An adequately functioning outside social network is considered important to support the patient during treatment and during his reintegration into society. Treatment progress is evaluated every three months by the treatment team, which includes the supervising psychologist, group leaders, the psychotherapist and the social worker of the patient.

The treatment model of the Van der Hoeven hospital is eclectic. A treatment program is offered, composed of individual and/or group therapy; job training, education, creative arts, and sports. Patients participate in group therapy programs, such as social skills training, aggression management, and substance abuse treatment. There are special group programs for pedophilic and adult-victim sexual offenders. Almost all patients receive individual cognitive-behavioral therapy, with an emphasis on diminishing violence risk through interventions aimed at increasing the patient's insight and control over his behavior. The cognitive-behavioral therapy integrates several approaches, such as Young's (1994) schema-focused therapy for personality disorders, Linehan's (1993) dialectical behavior therapy and the offense script and relapse prevention method (Laws, 1989).

3.2. Participants

Subjects were 87 male forensic psychiatric patients admitted to the hospital between January 1, 1996 and May 31, 2001, who consented to administration of the baseline (T1) and follow-up (T2) assessments, which were part of routine outcome monitoring procedures at the time. The sample represents approximately 70% of available male subjects admitted to the hospital in the above-mentioned period. Patients participated on a voluntary basis, and the remainder were either not examined or provided incomplete data (e.g., no T1 or T2 data), as a result of refusal of an interview, early referral to another facility, or severe

clinical symptoms. In fact, 28 patients were tested only at T1 and were therefore excluded from the present study. The mean length of time between baseline assessment (T1) and retest (T2) was 620 days or 20 months ($SD = 4$ months), varying from 15 to 36 months (Median = 19 months). Due to various reasons (e.g., the functioning of the patient being too unstable, temporary refusal of the patient to cooperate), it was not possible to completely standardize the time between T1 and T2.

Mean age at admission was 30.1 years ($SD = 7$ years; range = 19 to 47). Most (77%) of the men were Caucasian, and the others were of Surinamese/Dutch Antillean (14%), Moroccan-Turkish (7%) or other descent (2%). Sixty patients (69%) had never been married nor lived in a common law marriage. Forty-nine percent of the sample was convicted for (attempted) murder/homicide, 25% for sexual offenses (i.e., sexual assault, rape, and child sexual abuse), and almost 20% for violent offenses (assault, robbery with violence, and threat); the others for crimes against property and arson (for baseline characteristics, see Table 1).

4. Procedure

As a standard procedure, in order to obtain insight into factors related to the patient's (sexual) violence risk and to develop an individual case formulation that provides a basic treatment plan and for monitoring the patient's progress during treatment, patients are administered a standardized battery of psychological evaluation instruments upon admission (T1). *Multimethod* assessments (i.e., semi-structured interviews, self-report inventories, performance based personality tests, staff ratings) are employed, because distinct assessment methods provide unique sources of data. On the basis of a large array of evidence, Meyer et al. (2001; see also Meyer, 1997) argue that "optimal knowledge in clinical practice (as in research) is obtained from the sophisticated integration of information derived from a multimethod assessment battery" (p. 155). Multimethod assessment is particularly important in forensic subjects, who tend to be more prone to defensive responding, faking good or faking bad than subjects who are not assessed under mandatory conditions. By using multiple methods to assess the same variables (e.g., impulsivity) the findings

with one method can be cross validated against the findings obtained with another method.

To provide information on treatment progress, as a standard procedure, patients were re-tested with the same assessment battery (interviews not included), within 18–24 months after admission (T2). All instruments were administered according to standard administration procedures explained in the respective test manuals.

All psychological assessments were conducted by a pool of 10 clinical psychologists with master's or doctoral degrees. All were experienced in assessment and/or treatment of forensic psychiatric patients (cf. Hildebrand & de Ruiter, 2004; Hildebrand, de Ruiter, de Vogel, & van der Wolf, 2002). The psychologists ($n = 5$) who administered the Rorschach Inkblot Method (RIM) were trained in the *Comprehensive System* (CS; see Section 5.3.1).³

5. Measures

5.1. Semi-structured interviews

5.1.1. Psychopathy Checklist – Revised

Upon admission, patients were interviewed to obtain data on PCL-R psychopathy.⁴ Psychopathy was assessed using the Dutch language version (Vertommen, Verheul, de Ruiter, & Hildebrand, 2002) of the PCL-R (Hare, 1991). The standard procedure for scoring the PCL-R – coding on the basis of a semi-structured interview and extensive file information – was employed in 77 of the 87 cases. Items were scored on a 3-point scale (0 = item does not apply, 1 = item applies to a certain extent, 2 = item definitely applies). PCL-R interviews were videotaped, for which patients had to give their written informed consent, and PCL-R ratings were made by (at least) two independent raters.⁵ The remaining 10 cases were coded by two independent raters, based on file information only.⁶ To optimize scoring accuracy (Hare, 1991, 1998), PCL-R consensus scores were used in all subsequent data-analyses. It should be noted that PCL-R scores for every patient in this study were established by at least one rater who participated in the interrater reliability study of the Dutch language version of the PCL-R (Hildebrand et al., 2002). The interrater reliability was excellent. The intraclass correlation coefficient (ICC), using a two-way random effects model, for the PCL-R total score was .88 for a single rater (Factor 1 = .76; Factor 2 = .83). Ratings were also internally consistent (Cronbach's alpha = .87 for the PCL-R total score).

The mean total PCL-R score (adjusted sum) for the entire sample of 87 patients was 20.97 ($SD = 8.37$), with a range from 3 to 38, a median score of 22 and a mode of 17. The kurtosis of the PCL-R total score was $-.716$ ($SE = .511$). PCL-R scores were normally distributed (Kolmogorov-Smirnov $Z = .625$, $p = .829$). The mean Factor 1 score was 9.19 ($SD = 3.88$), the mean Factor 2 score was 9.35 ($SD = 5.53$). Using a cut-off score of 26, which is often used in European research (e.g., Cooke, 1995; Grann et al., 1998; Hildebrand, de Ruiter, & de Vogel, 2004; Hildebrand, de Ruiter, & Nijman, 2004), 27 patients (31%) were classified as 'psychopaths'. Nine patients (10%) had a PCL-R score < 10.

Table 1

Sample characteristics of psychopathic (PCL-R \geq 22) and non-psychopathic (PCL-R<22) patients.

Characteristics	Sample ($n = 87$)	PCL-R<22 ($n = 44$)	PCL-R>22 ($n = 43$)	<i>p</i> -Value
Age (M \pm SD)	30.1 (7.0)	30.0 (6.6)	30.1 (7.5)	.975
Partner: no (%)	60 (69.0)	29 (65.9)	31 (72.1)	.533
Caucasian: yes (%)	67 (77.0)	35 (79.5)	32 (74.4)	.570
Conviction for murder/ homicide: yes (%)	43 (49.4)	24 (54.5)	19 (44.2)	.334
Conviction for sexual offense: yes (%)	22 (25.3)	16 (36.4)	6 (14.0)	.016
Conviction for violent offense: yes (%)	17 (19.5)	2 (4.5)	15 (34.9)	.000
Time to follow up in days (M \pm SD)	620 (109.8)	623 (127.4)	617 (89.6)	.795
Schizophrenia/other psychotic disorder: yes (%)	11 (12.6)	8 (18.2)	3 (7.0)	.116
Alcohol and/or drug abuse/ dependence: yes (%)	46 (52.9)	16 (36.4)	30 (69.8)	.002
Any PD ^a : yes (%)	70 (81.4)	30 (69.8)	40 (93.0)	.006
Number of PD diagnoses, if any (M \pm SD)	1.6 (1.2)	1.2 (1.1)	2.0 (1.2)	.002
Paranoid PD: yes (%)	16 (18.6)	5 (11.6)	11 (25.6)	.096
Antisocial PD: yes (%)	41 (47.7)	9 (20.9)	32 (74.4)	.000
Borderline PD: yes (%)	19 (22.1)	6 (14.0)	13 (30.2)	.069
Narcissistic PD: yes (%)	23 (26.7)	7 (16.3)	16 (37.2)	.028

Note. PCL-R = Psychopathy Checklist – Revised. PD = Personality Disorder.

^a With regard to PD diagnoses, valid data were obtained for 86 patients.

³ One of them (the second author) is active as Rorschach instructor.

⁴ Because PCL-R psychopathy assessment was not implemented until November 1997, 34 patients were administered the PCL-R interview at T2.

⁵ Twenty-four patients refused to give consent for videotaping the interview; eight of them agreed with a joint interview approach (one rater conducted the interview while a second rater was present as an observer); six refused the presence of a second observer, and PCL-R scores had to be based on the judgment of a single interviewer (MH or CdR).

⁶ Previously, we reported that comparisons between real-life interview and videotaped interview indicated that the information source (interview versus video) did not influence the raters' coding (Hildebrand et al., 2002). Hildebrand, de Ruiter, and de Vogel (2004) and Hildebrand et al. (2002) reported high interrater reliabilities for the Dutch PCL-R, also for file-only ratings.

5.1.2. Structured Interview for DSM-IV(-TR) Disorders of Personality

Also upon admission, Axis II diagnoses were obtained by administration of the Dutch translation of the Structured Interview for DSM-IV Disorders of Personality (SIDP-IV; de Jong, Derks, van Oel, & Rinne, 1996; Pfohl, Blum, & Zimmerman, 1994). Diagnostic criteria were rated after the interview was completed and the interviewer had also examined available file information. Unfortunately, no interrater reliability data were collected for Axis II diagnoses, but in most cases the scoring was reviewed by a second, senior-level clinical psychologist, who also knew the patient.⁷

5.2. Self-report inventories

In addition to the interview methods, patients completed a number of self-report measures, including the Minnesota Multiphasic Personality Inventory-2, the Interpersonal Checklist – Revised, and the Barratt Impulsiveness Scale.

5.2.1. Minnesota Multiphasic Personality Inventory-2 (MMPI-2)

The MMPI-2 (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989; Dutch translation: Derksen, de Mey, Sloore, & Hellenbosch, 1993) is a 567-item (true/false) personality inventory that provides information on a subject's personality and psychopathology. In addition to the primary three validity scales and 10 clinical scales, the MMPI-2 includes (Harris–Lingoes) subscales that identify distinct components within the more heterogeneous clinical scales as well as a series of supplementary and content scales that further identify more specific psychological symptoms. Interpretation of its scores is based on a vast body of empirical research, and the MMPI-2 has become one of the most widely used clinical personality inventories. Of special appeal to forensic psychologists is that the MMPI-2 includes various validity scales designed to detect deviant test-taking attitudes (e.g., defensiveness, malingering). The Dutch version of the MMPI-2 has its own representative normative database. These Dutch norms correspond highly with the norms of the original version of the MMPI-2 (e.g., Butcher, Derksen, Sloore, & Sirigatti, 2003; Sloore, Derksen, de Mey, & Hellenbosch, 1996). Mean internal consistency coefficients varied from .63 to .85, and 18-week interval test-retest reliability for clinical and content scales varied from .43 to .91. Also, validity studies among different patient samples revealed promising results (e.g., Egger, de Mey, Derksen, & van der Staak, 2003a, 2003b).

Prior to data analysis, all MMPI-2 test protocols were screened for inconsistent responding and protocol invalidity. MMPI-2s were considered invalid if 30 or more items were omitted, True Response Inconsistency (TRIN) T-scores were >100, and Variable Response Inconsistency (VRIN) T-scores were >80. Using these exclusion criteria, three MMPI-2 protocols were eliminated.

5.2.2. Interpersonal Checklist – Revised (ICL-R)

The 160-item Dutch revised version of the ICL (LaForge & Suczek, 1955; Dutch version: ICL-R; de Jong, van den Brink, & Jansma, 2000) was used to map interpersonal styles. Each item (rated present/absent) is assigned to one of 10 dimensions. Scores on each dimension can range from 0 to 16, and scores between 4 and 12 are generally considered indicative of adaptive behavior. Within each dimension, items are designed to range from adaptive manifestations of a particular interpersonal behavior to extreme and maladaptive manifestations. Dimensions are labeled PA (managerial-autocratic), BC (competitive-exploitive), DE (aggressive-blunt), FG (distrustful-skeptical), nFnG (reserved-alloof), HI (modest-self-effacing), JK (docile-dependent), LM (cooperative-

overconventional), NO (responsible-overgenerous), and nNnO (extra-vert-gregarious). The scores on these dimensions can be transformed into a vector score, which roughly describes the interpersonal style in terms of the degree of power or control in an interaction (dominance versus submissiveness) and the degree of affiliation (hostility versus friendliness/nurturance). The ICL-R manual provides information on the psychometric properties of the ICL-R in Dutch samples. The intercorrelations and the circumplex analysis support the hypothesized circular arrangement of the interpersonal styles. The test-retest reliability of the interpersonal styles is moderate to good (ICC's range from .57 to .83; de Jong et al., 2000).

5.2.3. Barratt Impulsiveness Scale-11 (BIS-11)

The BIS-11 (Patton, Stanford, & Barratt, 1995; see also Barratt, 1994) is one of the most commonly applied psychometric measures of impulsivity (Spinella, 2007). The scale consists of 30-items, all rated from 1 (rarely/never) to 4 (almost always/always). Total scores range from 30 to 120. The BIS-11 assesses impulsivity in three domains: an ideomotor factor (acting without thinking); a careful planning dimension (attention to details), and a coping stability factor (lack of concern for the future). A high score on one of the subdomains suggests a tendency to impulsiveness in that domain. The internal consistency of the BIS-11 is acceptable, with Cronbach's α around .80 for the total score (Carrillo-de-la-Pena, Otero, & Romero, 1993; Fossati, Di Ceglie, Acquarini, & Barratt, 2001; Patton et al., 1995). The BIS-11 has been used to distinguish between violent and non-violent parolees (Cherek, Moeller, Dougherty, & Rhoades, 1997). The BIS is also a good predictor of aggression. Wang and Diamond (1999), for example, using structural equation modeling to predict institutional aggression among 385 male mentally disordered offenders, found that BIS impulsivity was a good predictor of institutional aggression.

5.3. Performance-based personality test

5.3.1. Rorschach Inkblot Method (RIM)

Patients were administered the RIM (Rorschach, 1921/1942) using Exner's (2001) *Comprehensive System* (CS) scoring system. The RIM consists of the consecutive presentation of a set of 10 achromatic and chromatic inkblots published by Huber Verlag. The basic assumption underlying the Rorschach is that personality characteristics influence the response process that takes place after the examiner has presented each card and asked the subject "What might this be?" According to Exner (1993), the RIM is a problem-solving task, in which the stimulus features of the blots are just as relevant as the idiosyncratic projections of the subject. All Rorschach protocols were administered and scored by the administrator following standardized procedures outlined by Exner (1993, 2001), and independently rescored by another colleague. All 10 inkblots were presented in their fixed order without limitations to the number of responses, while adhering to the break-off rules for too many responses to the first cards. All Rorschach responses, including the inquiry, were type-written verbatim. Consensus ratings were obtained, which were used in all subsequent analyses. Interrater reliability was examined using 52 protocols. Interscorer reliability measures were obtained for eight major categories (i.e., location, pairs, Z-score, populars, form quality, contents, determinants, and special scores) of Rorschach CS variables, as recommended by Exner, Kinder, and Curtiss (1995). According to Exner et al. (1995), the expected percentage of agreement will vary to some extent by segments. Typically, the percentage of agreement for location, pairs, Z-score and populars should approach 100%; form quality and contents should be around 85%; determinants and special scores are likely to be lowest but should not fall below 80%. The percentage of agreement between raters in the present study was as follows for each major category of Rorschach CS variables: location and developmental quality = 92%; determinants = 88%; form quality = 84%; pairs = 95%; contents = 94%; populars = 97%; organizational

⁷ Note that nine of the patients were diagnosed using the Dutch translation (van den Brink & de Jong, 1992) of the Structured Interview for DSM-III-R (SIDP-R; Pfohl, Blum, Zimmerman, & Stangl, 1989). This is a consequence of the duration of the data collection, which started before the Dutch SIDP-IV became available.

activity (Z -score) = 92%; and special scores = 82%, indicating good to excellent interscorer reliability. In fact, agreement measures met the standards recommended by Exner et al. (1995). In accordance with Exner's (1993) requirements, only Rorschach protocols with a number of responses ≥ 14 were included in the study. Using this exclusion criterion, protocols (either T1 and/or T2) of 29 patients had to be excluded.

5.4. Staff ratings

To determine the degree of convergence between patient and observer ratings on the ICL-R, as a routine procedure in our hospital, the ICL-R was also completed by staff members (i.e., a consensus rating of at least two group leaders who interact with the patient on a daily basis).

5.5. File reviews

Lifetime Axis I diagnoses were established by the first author using all available data (e.g., previous psychological and psychiatric reports, prior diagnoses, current psychiatric or psychological assessments). No interrater reliability data were collected for Axis I diagnoses. However, for 76 patients (87%), diagnoses were reviewed by three independent raters (cf. Hildebrand & de Ruiter, 2004): a senior clinical psychologist and a senior psychotherapist of the hospital staff, and the second author. Missing diagnoses were added and disagreements between the four raters were discussed and resolved, and a set of final consensus diagnoses for the 76 patients in the sample was established. This procedure using consensus diagnoses was chosen to maximize diagnostic accuracy.

Furthermore, computerized hospital records were reviewed to provide information about treatment compliance, number of missed and attended psychotherapy sessions, work and educational activities, creative arts, and sports.

6. Missing data

Although 87 patients participated in the study, unfortunately, only the analyses involving the ICL-R self-report have an n of 87. With regard to the MMPI-2, three protocols were eliminated because they represented unacceptable response inconsistencies. Therefore, all analysis including MMPI-2 variables relate to $n = 84$. Furthermore, because the data-collection of BIS-11 and ICL-R staff ratings, as a routine procedure in the hospital, started approximately one year after the initial start of the study, analyses including BIS-11 and ICL-R staff ratings relate to $n = 66$. With regard to the RIM CS, protocols (either T1 and/or T2) of only 58 patients fulfilled the inclusion criterion of ≥ 14 responses. Forty-two patients have complete data (i.e., MMPI-2, ICL-R self-report, BIS-11, ICL-R staff ratings, and RIM CS data).

Due to many patients missing at least one assessment instrument, the variable sample size poses challenges for data-analyses. Therefore, each hypothesis will be tested once with all available data (variable sample sizes), and again with the smaller "constant sample" of those having complete data ($n = 42$).

For purpose of the analyses, psychopathic and non-psychopathic offenders were identified by performing a median split on the PCL-R (median = 22), with 44 patients classified as "non-psychopaths" (PCL-R < 22) and 43 patients classified as "psychopaths" (PCL-R \geq 22). For the constant sample, 22 patients were classified as non-psychopathic (PCL-R < 19) and 20 patients as psychopathic, based on a median split on the PCL-R (median = 19 for the constant sample).

7. Selected indicators of the dynamic risk factors

Variables were selected from the self-report inventories and the RIM as indicators of dynamic risk factors in order to examine change during treatment. Dynamic risk factors (i.e., egocentrism/narcissism,

hostility, impulsivity, lack of insight, and negative attitudes) were selected because they have support in the scientific literature as risk markers of violence. For the RIM CS variables, each variable has a clinically significant cut-off score, above which the score is considered to be problematic⁸; all scores were dichotomized (0 = normal; 1 = problematic). The frequencies (percentage scores) of problematic scores on each of the RIM CS variables were assessed for psychopathic and nonpsychopathic patients. The following 22 variables were selected as indicators of dynamic risk factors.

7.1. Egocentrism

7.1.1. MMPI-2

Two MMPI-2 scales were used as indicators of egocentrism: the Harris-Lingoes (H-L) Subscale Social Imperturbability (Pd3), and the H-L Subscale Ego Inflation (Ma4). Pd3 reflects a highly aggressive and insouciant sociability consistent with a desire to use interpersonal relationships to manipulate, intimidate, and exploit. Ma4 measures the grandiosity typical of the (hypo)manic person. High scorers usually have unrealistic appraisals of their own abilities and self-worth, and become angry when their importance is not appreciated.

7.1.2. RIM CS

$Fr + rF > 0$ (narcissistic-like tendencies) and $PER > 3$ (overly defensive of one's self-image; argumentative) were selected as indicators of egocentrism/narcissism. Gacono and Meloy (1994) reported that psychopaths, as a group, showed elevated narcissistic-like self-interest, i.e., elevated $Fr + rF$ and PER responses.

7.2. Hostility

7.2.1. MMPI-2

Selected variables were the *Anger Content scale (ANG)* and the *Overcontrolled-Hostility Supplementary scale (O-H)*. ANG is concerned with poorly controlled anger (Schill & Wang, 1990). The O-H scale is a hostility measure shown to be associated with aggressive and violent acts in correctional settings (Graham, 2011).

7.2.2. ICL-R

BC (competitive-exploitive) and DE (aggressive-blunt) scale scores were used as ICL-R hostility variables.

7.2.3. Staff ratings ICL-R

BC and DE scale scores were used as ICL-R hostility variables.

7.3. Impulsivity

7.3.1. MMPI-2

The clinical scale *Hypomania (Ma, scale 9)* and the H-L Subscale *Lack of Ego Mastery/Defective Inhibition (Sc5)* were used as MMPI-2 indicators of impulsivity. Ma is related to energy level, irritability, and egotism. High scorers' activities may become fragmented and scattered due to very high energy levels (Friedman, Lewak, Nichols, & Webb, 2001). A high score on Sc5 is related to not feeling in control of emotions or impulses, irritability, hyperactivity, and dissociation of affect.

7.3.2. BIS-11

The total score on the BIS-11 was used.

7.3.3. RIM CS

$CF + C > FC + 1$ (lack of emotional control) was used to measure impulsivity.

⁸ Guidelines for interpreting RIM CS variables are given by Exner (1993, 2001) and Weiner (2003).

7.4. Lack of insight

7.4.1. ICL-R

Perceived differences between patient and staff ratings on ICL-R dimensions *BC*, *DE*, and *FG* were considered indicators of lack of insight.

7.4.2. RIM CS

Lack of introspective capacity was measured using *FD = 0* (Exner, 1993).

7.5. Negative distrustful attitudes

7.5.1. MMPI-2

The Content scales *Cynicism (CYN)* and *Negative Treatment Indicators (TRT)* were used as indicators of a negative attitude. *CYN* measures to what extent an individual distrusts others because they act only out of self-interest. High scores on *TRT* are indicative for negative attitudes toward doctors and mental health treatment, being uncomfortable in self-disclosing and avoiding responsibility for own actions (e.g., Pope, Butcher, & Seelen, 2006).

7.5.2. ICL-R

The *FG* (distrustful-skeptical) scale score was chosen as indicator of distrustfulness.

7.5.3. Staff ratings ICL-R

The *FG* scale score was used as ICL-R distrustfulness variable.

7.6. Treatment compliance

Computerized hospital case files were reviewed for the number of missed and attended therapeutic activities. The attendance rate of (individual) psychotherapy sessions, work, educational activities, creative arts, and sports were considered indicators of treatment compliance.

8. Data analysis

Descriptive analyses were applied to determine the prevalence rate of several characteristics of the sample (e.g., Axis I and Axis II disorders). Differences in baseline characteristics between the two groups under study were tested using analysis of variance (ANOVA) or the chi-square test to determine the differences in frequency distribution between diagnostic categories.

In order to examine differences between the scores on the self report (i.e., MMPI-2, ICL-R, and BIS-11) and staff (ICL-R) outcome measures for both groups at baseline assessment (T1), ANOVA's were calculated. Additionally, the effect size Cohen's *d* (Cohen, 1988), which provides an indication of the magnitude of the difference between two sample means in relation to the distribution of the scores within the samples, was computed. Typically, an effect size of .20 is considered a small effect, .50 a moderate effect, and .80 is considered a large effect (Cohen, 1988). The comparison of differences on RIM CS variables for psychopathic and non-psychopathic offenders at T1 was carried out by using the *z*-test for proportions (frequency comparison).

Assessing treatment progress on the self-report and staff outcome measures was investigated using the General Linear Model (GLM) Repeated Measures (M)ANOVA procedure. Group (psychopathic versus non-psychopathic patients) was used as between-subjects factor and time (T1, T2) as within-subjects factor. Because of the missing data (most patients had missing data on at least one outcome measure), for the different dynamic risk factors, measures involving the same individuals were combined in the analyses. For RIM CS variables, the McNemar test was used to test the difference between T1 and T2 scores (paired proportions) for psychopathic and non-psychopathic patients, separately.

The relation between PCL-R scores and attendance rate (i.e., number of attended activities – number of planned activities \times 100) of individual psychotherapy sessions, work, education, sports, and creative arts was evaluated with Pearson product-moment correlations, with α set at .05.

9. Results

9.1. Diagnostic characteristics at baseline

Table 1 summarizes diagnostic characteristics of the sample. Eleven patients (13%) met criteria for schizophrenia or another psychotic disorder. Forty-six patients (52%) met criteria for (at least) one substance related disorder, i.e., any alcohol or other substance related disorder (Table 1). The prevalence of personality disorders in the sample was substantial; 70 patients (81%; valid data obtained for 86 patients) received at least one Axis II diagnosis (Table 1). Co-morbidity on Axis II was common: of the 70 patients given a diagnosis, 40 (57%) received multiple diagnoses. The mean number of personality disorders per patient, for patients with at least one disorder, was 1.93. As expected in forensic settings, most personality disorders were Cluster B disorders, followed by Cluster A. The most frequently diagnosed disorder was antisocial personality disorder ($n = 41$), followed by narcissistic ($n = 23$), borderline ($n = 19$), and paranoid personality disorder ($n = 16$).

The analyses further showed that 36% of the non-psychopathic offenders and 14% of the psychopathic offenders were convicted for a sexual offense. The difference is significant, $\chi^2 = 5.78$, $p = .016$, $\phi = -.26$. Psychopathic offenders, on the other hand, were significantly more likely to have been convicted for a violent offense, 35% versus 5%, $\chi^2 = 12.73$, $p < .001$, $\phi = .38$. With regard to Axis I and Axis II diagnoses, psychopathic offenders were more likely to be diagnosed with alcohol and/or drug abuse/dependence, $\chi^2 = 9.74$, $p = .002$, $\phi = .34$, any Axis II disorder, $\chi^2 = 7.68$, $p = .006$, $\phi = .30$, as well as antisocial $\chi^2 = 24.66$, $p < .001$, $\phi = .54$ and narcissistic, $\chi^2 = 4.81$, $p = .028$, $\phi = .24$, personality disorder. Also, the total number of personality disorder diagnoses was significantly greater for psychopathic patients, $M = 2.0$ versus $M = 1.2$, $F(1, 84) = 9.81$, $p = .002$, $d = .69$.

9.2. Baseline (T1) scores on indicators of egocentrism, hostility, impulsivity, lack of insight, and negative distrustful attitudes

Table 2 summarizes the baseline and retest scores for the 22 indicators of the five dynamic risk factors. Scores at the two time points are presented for the total sample, as well as for psychopathic (PCL-R ≥ 22) and non-psychopathic (PCL-R < 22) patients separately. Overall, psychopathic and non-psychopathic patients showed similar levels of disturbance at baseline assessment, although tests of between group differences on the self-report and staff indicators of dynamic risk factors show that the psychopathic group had significantly higher scores on the Egocentrism indicator *MMPI-2 Pd3*, $F(1, 82) = 12.43$, $p = .001$, $d = .77$, the Hostility indicators self-report *ICL-R BC* (competitive-exploitive), $F(1, 85) = 6.42$, $p = .013$, $d = .54$, self-report *ICL-R DE* (aggressive-blunt), $F(1, 85) = 5.05$, $p = .027$, $d = .47$, staff ratings of *ICL-R BC*, $F(1, 64) = 4.14$, $p = .046$, $d = .52$, and on the Impulsivity indicator *MMPI-2 Ma*, $F(1, 82) = 6.33$, $p = .013$, $d = .56$, at baseline. With regard to the RIM CS variables, *z*-tests for proportions indicated no significant differences between psychopathic and non-psychopathic patients at T1 (*z* varying from .15 to 1.31).

9.3. Treatment progress

Effect size analyses using Cohen's *d* revealed negligible change on the self-report and staff indicators of egocentrism, hostility, impulsivity, lack of insight, and negative attitudes for both groups, except for

Table 2Baseline (T1) and retest (T2) Scores on indicators of dynamic risk factors for psychopathic (PCL-R \geq 22) and non-psychopathic (PCL-R $<$ 22) patients.

Outcome criteria	N	Total sample			PCL-R $<$ 22			PCL-R \geq 22		
		Baseline (T1)	Retest (T2)	Effect size (Cohen's <i>d</i>)	Baseline (T1)	Retest (T2)	Effect size (Cohen's <i>d</i>)	Baseline (T1)	Retest (T2)	Effect size (Cohen's <i>d</i>)
<i>Egocentrism</i>										
MMPI-2 (M \pm SD)										
Pd3	84	50.9 (10.2)	53.3 (10.0)	-.24	47.2 (11.5)	50.2 (11.1)	-.26	54.6 (7.3)	56.3 (7.8)	-.23
Ma4	84	53.5 (11.0)	51.1 (10.8)	.22	53.4 (10.7)	50.7 (11.6)	.24	53.5 (11.4)	51.4 (10.0)	.14
RIM CS (%)										
Fr + rF > 0	58	26%	29%		27%	29%		25%	29%	
PER > 3	58	16%	9%		18%	6%		13%	13%	
<i>Hostility</i>										
MMPI-2 (M \pm SD)										
ANG	84	53.0 (13.2)	54.3 (12.3)	-.10	51.3 (12.9)	52.0 (9.9)	-.06	54.7 (13.4)	56.6 (14.0)	-.14
O-H	84	60.3 (12.3)	60.6 (12.3)	-.02	59.6 (11.3)	59.3 (11.1)	.03	61.1 (12.9)	61.9 (13.4)	-.06
ICL-R (M \pm SD)										
BC self-report	87	7.4 (2.5)	8.0 (2.2)	-.26	6.8 (2.4)	7.3 (2.4)	-.21	8.1 (2.4)	8.6 (1.7)	-.24
DE self-report	87	7.7 (2.4)	8.2 (2.2)	-.22	7.2 (2.6)	7.6 (2.2)	-.17	8.3 (2.0)	8.8 (2.0)	-.25
Staff ratings BC	66	7.2 (3.8)	7.0 (3.2)	.06	6.3 (4.0)	6.8 (3.1)	-.14	8.2 (3.3)	7.7 (3.2)	.15
Staff ratings DE	66	6.2 (3.1)	7.3 (2.4)	-.40	5.5 (3.0)	7.0 (2.4)	-.55	7.0 (3.0)	7.7 (2.4)	-.32
<i>Impulsivity</i>										
MMPI-2 (M \pm SD)										
Ma (scale 9)	84	58.9 (11.7)	56.4 (11.9)	.21	55.8 (10.3)	54.6 (11.1)	.11	62.1 (12.3)	58.3 (12.5)	.31
Sc5	84	55.1 (12.7)	55.1 (13.0)	.00	55.6 (13.3)	55.7 (14.1)	.01	54.5 (12.1)	54.4 (11.9)	.01
BIS-11 (M \pm SD)										
Total score	66	61.3 (8.3)	60.3 (7.2)	.13	61.2 (8.0)	59.6 (6.3)	.22	61.3 (8.7)	60.9 (8.0)	.05
RIM CS (%)										
CF + C > FC + 1	58	33%	36%		41%	44%		21%	25%	
<i>Lack of Insight</i>										
ICL-R (M \pm SD)										
BC (self-staff)	66	0.3 (3.7)	0.8 (3.2)	-.14	0.4 (4.1)	0.5 (3.3)	-.03	1.1 (3.1)	1.5 (3.0)	-.13
DE (self-staff)	66	1.7 (2.9)	1.0 (3.1)	.23	1.8 (2.9)	0.8 (2.9)	.34	1.5 (3.0)	1.1 (3.3)	.13
FG (self-staff)	66	-2.4 (3.8)	-3.0 (3.0)	.18	-2.2 (3.8)	-3.2 (3.2)	.28	-2.7 (3.9)	-3.2 (2.8)	.04
RIM CS (%)										
FD = 0	58	52%	57%		56%	50%		46%	67%	
<i>Negative attitudes</i>										
MMPI-2 (M \pm SD)										
CYN	84	49.0 (12.4)	45.9 (11.1)	.26	47.9 (13.0)	45.2 (10.6)	.23	50.1 (11.7)	46.6 (11.6)	.30
TRT	84	52.2 (12.2)	48.6 (10.3)	.32	54.5 (14.7)	48.9 (11.5)	.42	49.8 (8.7)	48.3 (9.1)	.17
ICL-R (M \pm SD)										
FG self-report	87	7.2 (2.8)	7.5 (2.8)	-.11	6.9 (2.7)	7.4 (2.5)	-.19	7.6 (2.9)	7.5 (3.1)	.03
Staff ratings FG	66	9.7 (3.3)	11.1 (2.4)	-.48	9.3 (3.6)	11.0 (2.5)	-.55	10.1 (2.9)	11.1 (2.2)	-.39

Note. PCL-R = Psychopathy Checklist – Revised. MMPI-2 = Minnesota Multiphasic Personality Inventory-2. Pd3 = Harris-Lingoes Subscale Social Imperturbability. Ma4 = Harris-Lingoes Subscale Ego Inflation. RIM CS = Rorschach Inkblot Method, Comprehensive System scoring. ANG = Anger Content scale. O-H = Overcontrolled-Hostility Supplementary scale. ICL-R = Interpersonal Checklist – Revised. BC = competitive-exploitive; DE = aggressive-blunt. Ma = Hypomania scale. Sc5 = H-L Subscale Lack of Ego Mastery/Defective Inhibition. BIS-11 = Barratt Impulsiveness Scale-11. FG = skeptical-distrustful. CYN = Cynicism Content scale. TRT = Negative Treatment Indicators Content Scale.

ICL-R DE staff ratings for low-psychopathic patients, Cohen's *d* = .55 (Table 2).

9.3.1. Egocentrism

9.3.1.1. *MMPI-2*. The repeated-measures MANOVA indicated that there was a significant main effect for *time*, $F(2, 81) = 3.83$, $p = .026$, as well as for *group*, $F(2, 81) = 6.97$, $p = .002$. No *group* \times *time* interaction effect was found, $F(2, 81) = .23$, $p = .793$. According to the tests of the between subjects effect, psychopathic and non-psychopathic patients differed significantly on *Pd3*, $F(1, 82) = 13.11$, $p = .001$. However, both the psychopathic and non-psychopathic patients became more aggressive and manipulative (*Pd3*; $F(1, 82) = 6.16$, $p = .015$).

9.3.1.2. *RIM CS*. McNemar tests for the total sample, as well as for the low and high psychopathy groups indicated that patients did not change significantly on the RIM CS indicators of egocentrism ($p > .289$).

9.3.2. Hostility

9.3.2.1. *MMPI-2*. There was no difference between the scores of psychopathic and non-psychopathic patients (no main effect for *group*, $F(2, 81) = 2.99$, $p = .056$). Scores on *ANG* and *O-H* for both the psychopathic and non-psychopathic patients remained the same (main effect for *time*, $F(2, 81) = 1.00$, $p = .373$). The *group* \times *time* interaction was not significant, $F(2, 81) = .45$, $p = .637$.

9.3.2.2. *ICL-R self-reports*. There were significant main effects for *time*, $F(2, 84) = 3.50$, $p = .035$, and for *group*, $F(2, 84) = 5.65$, $p = .005$. No *group* \times *time* interaction effect was found, $F(2, 84) = .004$, $p = .996$. The tests of the between subjects effect indicated that psychopathic and non-psychopathic patients differed significantly on *BC* and *DE*, $F(1, 85) = 9.51$, $p = .003$, and $F(1, 85) = 7.19$, $p = .007$, respectively. Both the psychopathic and non-psychopathic patients became more competitive-exploitive on *BC*, $F(1, 85) = 5.35$, $p = .023$, and more dominant and assertive on *DE*, $F(1, 85) = 4.11$, $p = .046$.

9.3.2.3. *Staff ratings ICL-R*. A significant main effect for time was found, $F(2, 63) = 5.41, p = .007$. According to staff, both psychopathic and non-psychopathic patients became more dominant and assertive on *DE*, $F(1, 64) = 7.97, p = .006$. The main effect for *group* was not significant, $F(2, 63) = 2.73, p = .102$, and the *group* × *time* interaction was also not significant, $F(2, 63) = .68, p = .509$.

9.3.3. Impulsivity

9.3.3.1. *MMPI-2*. A significant main effect for *group* was found, $F(2, 81) = 3.90, p = .024$, indicating that scores of psychopathic and non-psychopathic patients were different. The tests of the between subjects effects indicated that psychopathic patients scored higher on *Ma*, $F(1, 82) = 5.23, p = .025$. There was no main effect for *time*, $F(2, 81) = 2.17, p = .121$ or a significant *group* × *time* interaction, $F(2, 81) = .58, p = .563$.

9.3.3.2. *BIS-11*. On the *BIS-11* scores, there was no difference between the scores of psychopathic and non-psychopathic patients (no main effect for *group*, $F(1, 64) = .19, p = .667$). The *BIS-11* total score hardly changed during treatment for both the psychopathic and non-psychopathic patients (main effect for *time*, $F(1, 64) = 1.23, p = .273$). The interaction between *group* and *time* did not reach statistical significance, $F(1, 64) = .47, p = .496$.

9.3.3.3. *RIM CS*. No significant differences were found on *CF + C > FC + 1* (McNemar test, $p > .804$).

9.3.4. Lack of insight

9.3.4.1. *ICL-R (perceived differences between patient and staff)*. There were no main effects for *time*, $F(3, 61) = 1.99, p = .125$ and *group*, $F(3, 61) = .12, p = .946$, nor was there a significant *group* × *time* interaction effect, $F(3, 61) = .29, p = .832$.

9.3.4.2. *RIM CS*. No significant differences were found on *RIM CS FD = 0* (McNemar test, $p > .180$).

9.3.5. Negative distrustful attitudes

9.3.5.1. *MMPI-2*. A significant main effect was found for *time*, $F(2, 81) = 5.18, p = .008$. The main effect for *group* did not reach statistical significance, $F(2, 81) = 2.84, p = .065$. Also, there was no *group* × *time* interaction effect, $F(2, 81) = 1.71, p = .185$. At T2, both psychopathic and non-psychopathic patients were less cynical on *Cyn*, $F(1, 82) = 8.09, p = .006$, and had less negative attitudes toward mental health treatment on *TRT*, $F(1, 82) = 6.76, p = .011$.

9.3.5.2. *ICL-R self reports*. There was no difference between the scores of psychopathic and non-psychopathic patients on *FG* (no main effect for *group*, $F(1, 85) = .75, p = .390$). Scores on *FG* for both the psychopathic and non-psychopathic patients remained the same (main effect for *time*, $F(1, 85) = .51, p = .476$). The *group* × *time* interaction was not significant, $F(1, 85) = .86, p = .357$.

9.3.5.3. *Staff ratings ICL-R*. A significant main effect for *time* was found, $F(1, 64) = 8.60, p = .005$, indicating that psychopathic as well as non-psychopathic patients became more distrustful-skeptical on *FG*, according to staff. There was no main effect for *group*, $F(1, 64) = .68, p = .412$, and no *group* × *time* interaction effect, $F(1, 64) = .67, p = .417$.

9.4. Constant sample

9.4.1. Baseline scores constant sample

With regard to the smaller constant sample ($n = 42$) of those having complete data, similar results were found. Compared to non-

psychopathic patients, psychopathic patients had significantly higher scores on three of the 22 indicators: the *egocentrism* indicator *MMPI-2 Pd3* [$M = 53.4$ ($SD = 8.8$) versus $M = 45.3$ ($SD = 11.9$), $F(1, 40) = 6.19, p = .017, d = .77$], and the *hostility* indicators self-report *ICL-R BC* [$M = 9.3$ ($SD = 2.9$) versus $M = 6.8$ ($SD = 2.4$), $F(1, 40) = 8.82, p = .005, d = .93$] and self-report *ICL-R DE* [$M = 8.0$ ($SD = 2.5$) versus $M = 6.9$ ($SD = 2.9$), $F(1, 40) = 5.58, p = .023, d = .41$]. A trend was found indicating that psychopathic patients had significantly higher scores on the *impulsivity* indicator *MMPI-2 Ma* [$M = 63.3$ ($SD = 11.2$) versus $M = 56.7$ ($SD = 10.0$), $F(1, 40) = 3.97, p = .053, d = .62$]. Z-tests for proportions indicated no significant differences between both groups at baseline for the *RIM CS* variables (z varying from .015 to .771).

9.4.2. Treatment progress constant sample

With regard to treatment progress, in general, similar results were found for the constant sample as for the total sample. Below, only results of indicators with a significant effect are reported. For the self-report indicators of 'egocentrism', there were significant main effects for *time* and *group*, $F(2, 39) = 3.89, p = .029$, and $F(2, 39) = 5.85, p = .002$, respectively. No *group* × *time* interaction effect was found, $F(2, 39) = .60, p = .556$. The tests of the between subjects effect indicated that psychopathic and non-psychopathic patients differed significantly on *Pd3*, $F(1, 40) = 11.99, p = .001$. Both groups became more aggressive and manipulative over time on *Pd3*, $F(1, 40) = 6.66, p = .017$.

With regard to the self-report and staff measures of 'hostility', significant main effects were found for *time* and *group*, $F(6, 35) = 3.89, p = .004$, and $F(6, 35) = 2.82, p = .024$, respectively. Both groups differed significantly on (self report) *BC* and *DE*, $F(1, 40) = 10.37, p = .003$, and $F(1, 40) = 5.79, p = .021$, respectively. There was no significant *group* × *time* interaction effect, $F(6, 35) = .004, p = .996$. According to staff, psychopathic and non-psychopathic patients became more dominant and assertive on *DE*, $F(1, 40) = 10.05, p = .003$.

For the indicators of 'impulsivity' the main effect for *group* was significant, $F(3, 38) = 3.51, p = .024$. Tests of the between subjects effects indicated that *Ma* scores for psychopathic patients were higher $F(1, 40) = 4.14, p = .049$. No main effect for *time* was found, $F(3, 38) = 1.42, p = .251$, and there was no significant *group* × *time* interaction effect, $F(3, 38) = .40, p = .754$.

A significant main effect for *time* was found for the indicators of 'negative and distrustful attitudes', $F(4, 37) = 12.30, p < .001$. There was no main effect for *group*, $F(4, 37) = .82, p = .520$, and also no *group* × *time* interaction effect, $F(4, 37) = .11, p = .395$. Both psychopathic and non-psychopathic patients were less cynical on *Cyn* at T2, $F(1, 40) = 7.41, p = .010$, and less negative toward mental health treatment on *TRT*, $F(1, 40) = 6.37, p = .016$. Also, according to staff, both groups became more distrustful-skeptical during treatment on *FG*, $F(1, 40) = 5.56, p = .023$.

With regard to the *RIM CS* indicators of dynamic risk factors, no significant differences between T1 and T2 scores were found (McNemar, $p > .503$) in the constant sample.

9.5. Treatment compliance

The correlations between PCL-R scores and attendance rates of different therapeutic activities are displayed in Table 3.

The PCL-R total score, as well as PCL-R Facet 2, Facet 3 and Facet 4 scores were negatively associated with attendance at work and educational activities (r between $-.30$ and $-.36, p < .01$); the higher the PCL-R score, the worse the patient's attendance rate. Facet 4 was also negatively associated with creative arts' attendance. Attendance of individual psychotherapy sessions and sports was not related to PCL-R scores. Facet 1 was not in any way related to treatment compliance.

With regard to the correlations between PCL-R scores and attendance rates of different therapeutic activities for the constant sample,

Table 3
Correlations between psychopathy scores and attendance rate of therapeutic activities.^a

Therapeutic activities	PCL-R score													
	Total	Factor 1 (1st ed.)		Factor 2 (1st ed.)		Facet 1 (2nd ed.)		Facet 2 (2nd ed.)		Facet 3 (2nd ed.)		Facet 4 (2nd ed.)		
Psychotherapy ^b	.16	(-.13)	.19	(-.08)	.15	(.10)	.16	(-.04)	.07	(-.17)	.09	(.20)	-.02	(-.05)
Work ^c	-.33***	(-.36*)	-.16	(-.21)	-.34***	(-.36*)	-.14	(-.18)	-.11	(-.12)	-.30**	(-.36*)	-.36***	(-.35*)
Education ^c	-.34***	(-.42***)	-.13	(-.23)	-.32**	(-.44***)	-.16	(-.20)	-.06	(-.14)	-.34***	(-.42***)	-.35***	(-.47**)
Sports ^c	-.11	(-.23)	-.02	(-.11)	-.09	(-.19)	-.03	(-.07)	-.02	(-.07)	-.05	(-.17)	-.18	(-.30*)
Creative arts ^b	-.14	(-.27)	-.02	(-.14)	-.08	(-.27)	.07	(-.13)	.01	(-.03)	-.08	(-.27)	-.22*	(-.31*)

Note. PCL-R = Psychopathy Checklist – Revised. 1st ed. = First edition of the PCL-R (Hare, 1991). 2nd ed. = Second edition of the PCL-R (Hare, 2003).

^a Results for the constant sample ($n = 42$) are in brackets.

^b $N = 84$

^c $N = 86$.

* $p < .05$ (two-tailed).

** $p \leq .01$ (two-tailed).

*** $p \leq .001$ (two-tailed).

results are comparable with the results of the total sample (see Table 3).

9.6. Comparison between patients included in the sample and patients tested only upon admission

Due to various reasons (e.g., refusal, referral to another facility), 28 patients were tested only at T1 and were therefore excluded from the present study. It is relevant whether psychopathy scores and scores on dynamic risk factors for these patients were different from patients who provided T1 and T2 data sets. If so, there could be a selection bias. Comparisons between patients who did and who did not provide a data set at retest, indicated that the PCL-R total score of the dropouts was significantly higher ($M = 24.8$, $SD = 7.4$) than the mean PCL-R score of those who provided T1 and T2 data ($M = 21.0$, $SD = 8.4$), $t(1, 108) = 1.99$, $p = .048$. PCL-R Factor 1 and Factor 2 scores of both groups, however, did not differ significantly, $t(1, 107) = 1.57$, $p = .119$ and $t(1, 107) = 1.68$, $p = .095$, respectively. Also, dropouts and completers did not differ significantly in terms of the categorical diagnosis of psychopathy ($PCL-R \geq 26$), $\chi^2(1, 115) = 1.32$, $p = .250$. With regard to the 18 self-report, and staff outcome measures of dynamic risk factors, ANOVA's revealed significant differences on only two indicators. It was found that the dropouts were significantly more cynical on CYN, $F(1, 109) = 5.32$, $p = .023$ and more egocentric on Ma4, $F(1, 109) = 5.64$, $p = .019$, compared to patients who provided T1 and T2 data.

10. Discussion

The purpose of the present study was to explore the extent to which change in specific psychological measures of dynamic risk factors relevant to re-offending could differentiate psychopathic from non-psychopathic patients during inpatient forensic psychiatric treatment. It was hypothesized that upon admission to the hospital, psychopathic offenders (median split; $PCL-R \geq 22$) would show higher levels of disturbance on indicators of dynamic risk than non-psychopathic offenders. We expected the high PCL-R group to show more limited improvement after two years of treatment than patients with low PCL-R scores. In addition, this study investigated the relationship between psychopathy and treatment compliance, as indicated by the attendance rates of different kinds of therapeutic activities. We expected treatment noncompliance to be more characteristic of patients with high PCL-R scores.

Three major findings are evident from our study. First, overall, psychopathic and non-psychopathic patients showed similar levels of disturbance at baseline assessment, although tests of between group differences on the self-report and staff indicators of dynamic risk factors showed that the psychopathic group showed significantly higher levels of disturbance on five of the 22 indicators of dynamic

risk. At baseline, psychopathic patients had significantly higher scores on the Egocentrism indicator *MMPI-2 Pd3*, as well as on the Hostility indicators *ICL-R BC* (self report and staff ratings), *ICL-R DE* (self report), and on the Impulsivity indicator *MMPI-2 Ma*. Thus, our first hypothesis was not confirmed. Contrary to our expectation, psychopaths and non-psychopaths showed more similarities than differences in dynamic risk factors at baseline.

Second, the sample as a whole did not improve on most of the indicators of dynamic risk after 20 months of treatment. However, it was found that both the psychopathic and non-psychopathic patients became more interpersonally aggressive, more competitive-exploitive and more dominant and assertive during treatment. Furthermore, both groups were less cynical and less negative toward mental health treatment at T2, according to the *MMPI-2* measures. *ICL-R* staff ratings, on the other hand, indicated that psychopathic as well as non-psychopathic patients became more distrustful-skeptical during treatment. Contrary to our second hypothesis, no differential treatment response was found between psychopathic and nonpsychopathic patients. One might suggest that the lack of significantly different results between the two groups may be due, at least in part, to the fact that a relatively low cutoff score of 22 was used to divide patients into a psychopathic and non-psychopathic group. However, using cutoff scores of 26 and 30 also revealed that psychopathy was not related to treatment progress.

Third, our results provide partial support for the association between psychopathy and treatment noncompliance, i.e., the extent to which the patient actually participates in the assigned treatment program. High PCL-R total, Facet 2, Facet 3 and Facet 4 scores were significantly related to a lower level of involvement in education and work, suggesting that psychopathic patients tend to put less effort into the treatment program. In addition, high Facet 4 scores were also significantly associated with a lower level of involvement in creative arts. A potential reason for the differential prediction of compliance by activities (education/work versus psychotherapy/sports) might be that education and work are activities that require a certain effort and work towards long-term goals, and psychopathic patients in general are not very good at this. On the contrary, sports and psychotherapy are probably more directly rewarding for psychopathic patients, in so far they provide immediate attention (psychotherapy) and a fun way to spend time (sports).

Two possible explanations arise as to why changes (improvements) in dynamic risk factors were minimal during the 20 months treatment course. A first explanation for our findings is that the patients in our sample have serious personality pathology, which is very difficult to change; a second concerns the type of treatment program offered at the hospital. With regard to the first explanation, it should be mentioned that our sample consisted only of high risk mentally disordered offenders. Changes in dynamic risk factors such as egocentrism, anger, impulsivity, lack of insight, and negative

distrustful attitudes may be very hard to accomplish. It may very well be that high risk cases (psychopathic or non-psychopathic) should receive more intensive services over a longer period than 20 months to establish change. To illustrate this point, [Belfrage and Douglas \(2002\)](#) examined change in violence risk factors in forensic psychiatric patients across multiple assessment periods, using the HCR-20. They found that, in a sub-sample of 70 high risk patients undergoing long-term treatment, HCR-20 dynamic risk factors changed only modestly. Possibly, patients improve on indicators of dynamic risk factors at a slower pace. Most patients in our sample will continue treatment for an extended period and future assessments (e.g., after five years of treatment) may reveal the hypothesized changes. On the other hand, [Chakhssi, de Ruiter, and Bernstein \(2010\)](#), in a Dutch sample of 74 forensic psychiatric patients, comparable to the one used in this study, did find that both psychopathic and non-psychopathic patients changed with regard to level of insight according to the Behavioural Status Index (BEST-Index; [Reed, Woods, & Robinson, 2000](#)) over the course of 20 months. Maybe the BEST-Index is a better indicator for measuring improvements on insight than the indicators used in the present study, i.e., the perceived differences between patients and staff on three ICL-R segments.

The second possible explanation for our findings is that it may be that the treatment program provided at the hospital deserves review or alteration. Although the program takes into account that patients present a diversity of problems, it is largely based on the principles of the therapeutic community: an environment is created in which complex interpersonal and community processes become central therapeutic agents. The treatment program is eclectic and only partially focused on systematic change in the dynamic risk factors we studied. Previous research indicates that milieu therapies permit psychopathic patients to con staff into believing they are making progress ([Hare et al., 2000](#); [Hobson et al., 2000](#)). Alteration of the treatment program into a program based on the principles of risk, need and responsivity, focusing on reducing the risk of violence and destructiveness by modifying the cognitions and behaviors that directly precipitate violent behavior, may help maximize change. The standard of service delivery could also be increased by forming homogenous groups of patients allowing the development of specialized wards to target the needs of different groups of patients (e.g., [Rice et al., 1990](#)), with the specificity of each ward being based on both patient's treatment needs and security requirements. As a standard procedure, criminogenic needs identified during baseline assessment should become treatment targets, and for each an explicit treatment plan needs to specify how change is to be accomplished. Appropriate interventions delivered in this manner may produce favorable results in the treatment of this high risk group of offenders. In recent years, the Van der Hoeven Hospital has implemented a specific treatment program tailored to the risk, needs and responsivity of patients with high PCL-R scores ([Kroger et al., 2011](#)).

Our results provide support for the negative association between psychopathy and treatment compliance, suggesting that psychopathic patients tend to put less effort into the treatment program. Attendance rates of individual psychotherapy and sports, however, were not related to PCL-R scores. In general, these findings are in line with earlier findings (e.g., [Hobson et al., 2000](#); [Ogloff et al., 1990](#)) supporting the value of the PCL-R as a significant correlate of treatment compliance in forensic inpatients. For example, [Hobson et al. \(2000\)](#) found that the PCL-R total and Factor 1 score were associated with negative behaviors shown in therapy groups. They also reported significant correlations between PCL-R scores (total and factor scores) and behavior on the wing. Off wing activities (e.g., education, charity work) were inversely related to Factor 2 scores. Contrary to the research of [Hobson et al. \(2000\)](#), who found evidence that attention should be given equally to Factor 1 and Factor 2 scores, the present research indicates that Factor 2, in particular Facet 4, is of particular importance in forecasting and effectively managing treatment

compliance. In the present study, the Factor 1 score was not associated with treatment compliance. It should be noted though that comparing our findings with earlier studies is complicated by the fact that different studies use different criterion measures to operationalize treatment compliance. For example, [Hobson et al. \(2000\)](#) used an officer-completed behavior checklist to monitor institutional adjustment (e.g., "manipulates others"; "tells lies"; "inflated sense of self-importance"), whereas in the present study computerized hospital records were reviewed to provide information about the level of involvement in therapeutic activities. This might explain why in the [Hobson et al.](#) study the Factor 1 score was associated with poor adjustment to the therapeutic community.

A number of limitations of this study should be noted. The first and most significant limitation is the loss of subjects throughout the study. Patients could not be obliged to participate in the study, and some refused to cooperate at the time of retest; others could not be tested because their functioning was too unstable. Although 87 patients are involved, few analyses have an N of 87. Due to many patients missing at least one assessment instrument, the variable sample size posed challenges for data-analyses. Missing data might account for lower statistical power, smaller effect sizes, or error prone parameter estimates. Generally, lack of power may occur when the sample size is small and when the base rate of psychopathy is low. In the present study, 27 patients (31%) were diagnosed as psychopathic ($PCL-R \geq 26$). However, because of the rather small sample size and the base rate problem, psychopathic and non-psychopathic offenders were identified by performing a median split on the PCL-R and each hypothesis was tested once with all available data, and again with a smaller "constant sample" of those having complete data. In both samples, there were not even trends in the direction of the hypothesis that psychopathic patients show more limited improvement after two years of inpatient treatment than non-psychopathic offenders.

A second important limitation is the lack of an adequate no treatment control group. We did not use an experimental design, one in which treatment and control groups were formed prior to treatment using random assignment and in which group equivalence with respect to theoretically important variables was evaluated (e.g., [Hemphill & Hart, 2002](#)). Admission to the forensic hospital is based on specific legal criteria, and an equivalent control group could not be formed.

Third, data collection was restricted to one forensic psychiatric hospital. The extent to which these treatment outcome findings apply to the treatment programs offered in other (Dutch) forensic psychiatric hospitals is not known. The level of success may be different for other hospitals, which offer different programs. However, [Chakhssi et al. \(2010\)](#) also found no significant differences between psychopathic and non-psychopathic patients on risk-related behaviors (e.g., adaptive social behavior, attribution of responsibility, self-regulation strategies) during 20-months of inpatient forensic treatment.

Another limitation may be that the measures we employed were not sensitive enough to evaluate change in dynamic risk variables in a complex treatment regime, such as the one in our hospital ([Hughes et al., 1997](#)). However, the notable stability of Rorschach indicators in this study contrasts with earlier treatment outcome studies which found positive changes on (many of) these indicators ([Abraham, Lepisto, Lewis, Schultz, & Finelberg, 1994](#); [Exner & Andronikof-Sanglade, 1992](#); [Weiner & Exner, 1991](#)). [Weiner and Exner \(1991\)](#), for example, used 27 RIM CS indicators of adjustment difficulty to evaluate treatment effects in a sample of (non-forensic) outpatients, mainly suffering from anxiety disorders or depression. They reported significant improvement for the patients on indicators of stress tolerance, affect modulation, introspective abilities, quality of interpersonal relationships, and egocentricity. Improvement could be shown after short-term treatment but was even greater after longer-term therapy. Moreover, [Grønnerød \(2004\)](#) conducted a meta-analysis of studies using the Rorschach method to assess

personality changes following psychotherapy, and concluded that the method shows moderate sensitivity. According to the author, overall results were strong and show that the RIM seemed to perform as well as other instruments. Furthermore, researchers at the [Network for Addiction Treatment Novadic \(2003\)](#), investigating changes in the self-perceptions of interpersonal behavior in patients with substance abuse disorder reported that significant changes on the control vector of the ICL-R were accounted for by changes on competitive-exploitive, aggressive-blunt, distrustful-skeptical and docile-dependent dimensions, indicating that the ICL-R is sensitive to evaluate change.

A final limitation we would like to mention is the fact that the diagnostic procedures used to assess Axis I and Axis II pathology were not ideal. The use of a (semi-) structured interview for the assessment of lifetime Axis I diagnoses in combination with a record review would have been preferable. However, limited staff necessitated us to choose either an Axis I or Axis II semi-structured interview. Because patient files tended to include quite detailed information on Axis I pathology, we opted to employ the Axis II interview. Unfortunately, we were not able to assess interrater reliabilities for Axis I and Axis II diagnoses. However, in our opinion, with regard to Axis I and Axis II diagnoses, the diagnostic procedures were optimal for the study's clinical circumstance as they are probably more thorough than the usual clinical interview by a single professional. Further, although this information is important, it was not central to our study's purpose.

The present study requires replication, preferably including patients from different institutions, before more definitive conclusions can be drawn. Future researchers may wish to consider comparing specialized dynamic risk treatment programs with the 'treatment as usual' program in their particular hospital. Future research should also take specific treatment objectives and strategies into account to examine the differences in responsiveness between psychopathic and non-psychopathic patients.

The objective of inpatient forensic treatment is to effect changes in long-term patterns of maladaptive behavior in order to reduce violence risk. The overall trend over the course of therapy indicated that the sample as a whole did not improve on most of the indicators of dynamic risk factors. Psychopathy was also unrelated to treatment outcome. Since the treatment program was generally unsuccessful in addressing dynamic risk factors with either group of patients, no conclusions can be drawn about relative success, and this study thus has little to say about differential treatment-related change in psychopathic patients. What it *can* say is that other, non-psychopathic patients too, did not show much change over the course of treatment – yet a different treatment program might show different results. The results imply that a lack of progress should not be all too easily attributed to psychopathy, and psychopathic patients should not be excluded from forensic inpatient treatment (see also [Skeem, Monahan, & Mulvey, 2002](#)).

References

Abraham, P., Lepisto, L. B., Lewis, M. G., Schultz, L., & Finelberg, S. (1994). Changes in Rorschach variables of adolescents in residential treatment: An outcome study. *Journal of Personality Assessment*, 62, 505–514.

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.

Andrews, D. A., & Bonta, J. (1994). *The psychology of criminal conduct*. Cincinnati, OH: Anderson.

Andrews, D. A., & Bonta, J. (2003). *The psychology of criminal conduct* (3th ed.). Cincinnati, OH: Anderson.

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, 17, 19–52.

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, 28, 369–404.

Barratt, E. S. (1994). Impulsiveness and aggression. In J. Monahan, & H. Steadman (Eds.), *Violence and mental disorder: Developments in risk assessment* (pp. 61–79). Chicago: University of Chicago Press.

Baumeister, R. F., Bushman, B. J., & Campbell, W. K. (2000). Self-esteem, narcissism, and aggression: Does violence result from low self-esteem or from threatened egotism? *Current Directions in Psychological Science*, 9, 26–29.

Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Psychological Review*, 103, 5–33.

Belfrage, H., & Douglas, K. S. (2002). Treatment effects on forensic psychiatric patients measured with the HCR-20 violence risk assessment scheme. *International Journal of Forensic Mental Health*, 1, 25–36.

Blackburn, R. (2001). Treatment in high security hospitals. In C. R. Hollin (Ed.), *Handbook of offender assessment and treatment* (pp. 525–535). Chichester, UK: Wiley.

Blair, R. J. R. (2006). Subcortical brain systems in psychopathy: The amygdala and associated structures. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 296–312). New York: Guilford.

Bonta, J., Law, M., & Hanson, K. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. *Psychological Bulletin*, 123, 123–142.

Bushman, B. J., & Baumeister, R. F. (1998). Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: Does self-love or self-hate lead to violence? *Journal of Personality and Social Psychology*, 75, 219–229.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring*. Minneapolis: University of Minnesota Press.

Butcher, J. N., Derksen, J. J., Sloore, H., & Sirigatti, S. (2003). Objective personality assessment of people in diverse cultures: European adaptations of the MMPI-2. *Behaviour Research and Therapy*, 41, 819–840.

Carrillo-de-la-Pena, M. T., Otero, J. M., & Romero, E. (1993). Comparison among various methods of assessment of impulsiveness. *Perceptual and Motor Skills*, 77, 567–575.

Chakhssi, F., de Ruiter, C., & Bernstein, D. (2010). Change during forensic treatment in psychopathic versus nonpsychopathic offenders. *Journal of Forensic Psychiatry & Psychology*, <http://dx.doi.org/10.1080/14789949.2010.483283> First published on 27 July 2010.

Cherek, D. R., Moeller, F. G., Dougherty, D. M., & Rhoades, H. (1997). Studies of violent and nonviolent male parolees: II. Laboratory and psychometric measurements of impulsivity. *Biological Psychiatry*, 41, 523–529.

Cleckley, H. (1976). *The mask of sanity* (5th ed.). St. Louis, MO: Mosby.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cooke, D. J. (1995). Psychopathic disturbance in the Scottish prison population: The cross cultural generalizability of the Hare's Psychopathy Checklist. *Psychology, Crime and Law*, 2, 101–118.

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment*, 13, 171–188.

Cooke, D. J., & Philip, L. (2001). To treat or not to treat? An empirical perspective. In C. R. Hollin (Ed.), *Handbook of offender assessment and treatment* (pp. 17–34). Chichester: Wiley.

D'Silva, K., Duggan, C., & McCarthy, L. (2004). Does treatment really make psychopaths worse? A review of the evidence. *Journal of Personality Disorders*, 18, 163–177.

de Jong, C. A. J., Derks, F. C. H., van Oel, C. J., & Rinne, Th. (1996). *Gestructureerd Interview voor de DSM-IV Persoonlijkheidsstoornissen [Structured Interview for DSM-IV Personality Disorders]*. Sint Oedenrode, The Netherlands: Stichting Verslavingszorg Oost Brabant.

de Jong, C. A. J., van den Brink, W., & Jansma, A. (2000). *ICL-R: Handleiding bij de vernieuwde Nederlandse versie van de Interpersonal Checklist (ICL) [Manual of the revised Dutch version of the Interpersonal Checklist (ICL)]*. Sint Oedenrode, The Netherlands: Novadic.

de Ruiter, C., & Hildebrand, M. (2003). The dual nature of forensic psychiatric practice: Risk assessment and management under the Dutch TBS-order. In P. J. van Koppen, & S. D. Penrod (Eds.), *Adversarial versus inquisitorial justice: Psychological perspectives on criminal justice systems* (pp. 91–106). New York: Kluwer/Plenum.

Dempster, R. J., & Hart, S. D. (2002). The relative utility of fixed and variable risk factors in discriminating sexual recidivists and nonrecidivists. *Sexual Abuse: A Journal of Research and Treatment*, 14, 121–138.

Derksen, J., de Mey, H., Sloore, H., & Hellenbosch, G. (1993). *MMPI-2: handleiding bij afname, scoring en interpretatie [MMPI-2: Manual for administration, scoring, and interpretation]*. Nijmegen, The Netherlands: PEN Test Publisher.

Douglas, K. S., Ogloff, J. R. P., Nicholls, T. L., & Grant, I. (1999). Assessing risk for violence among psychiatric patients: The HCR-20 risk assessment scheme and the Psychopathy Checklist: Screening Version. *Journal of Consulting and Clinical Psychology*, 67, 917–930.

Egger, J. I. M., de Mey, H. R. A., Derksen, J. J. L., & van der Staak, C. P. F. (2003a). Cross-cultural replication of the Five-Factor Model and comparison of the NEO-PI-R and MMPI-2 PSY-5 scales in a Dutch psychiatry sample. *Psychological Assessment*, 15, 81–87.

Egger, J. I. M., de Mey, H. R. A., Derksen, J. J. L., & van der Staak, C. P. F. (2003b). MMPI-2 and MCMI-III scores among Dutch inpatient substance abusers: Assessing correspondence and cross-cultural equivalence. *Current Psychology*, 22, 117–124.

Exner, J. E., Jr. (1993). *The Rorschach: A comprehensive system* (Vol. 1). (3rd ed.). New York: Wiley.

Exner, J. E., Jr. (2001). *A Rorschach workbook for the comprehensive system*. Ashville, NC: Rorschach Workshops.

Exner, J. E., & Andronikof-Sanglade, A. (1992). Rorschach changes following brief and short-term therapy. *Journal of Personality Assessment*, 59, 59–71.

Exner, J. E., Jr., Kinder, B. N., & Curtiss, G. (1995). Reviewing basic design features. In J. E. Exner Jr. (Ed.), *Issues and methods in Rorschach research* (pp. 145–158). Mahwah, NJ: Lawrence Erlbaum.

Fossati, A., Di Ceglie, A., Acquarini, E., & Barratt, E. S. (2001). Psychometric properties of an Italian version of the Barratt Impulsiveness Scale-11 (BIS-11) in nonclinical subjects. *Journal of Clinical Psychology*, 57, 815–828.

Friedman, A. F., Lewak, R., Nichols, D. S., & Webb, J. T. (2001). *Psychological assessment with the MMPI-2*. Mahwah, NJ: Erlbaum.

Gacono, C. B., & Meloy, J. R. (1994). *The Rorschach assessment of aggressive and psychopathic personalities*. Hillsdale, NJ: Erlbaum.

- Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offenders: What works! *Criminology*, 34, 575–607.
- Graham, J. R. (2011). *MMPI-2: Assessing personality and psychopathology* (5th ed.). New York: Oxford University Press.
- 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.
- Grønnerød, C. (2004). Rorschach assessment of personality changes following psychotherapy: A meta-analytic review. *Journal of Personality Assessment*, 83, 256–276.
- Hanson, R. K. (1998). What do we know about sexual offender risk assessment. *Psychology, Public Policy, and Law*, 4, 50–72.
- Hare, R. D. (1991). *The Hare Psychopathy Checklist – Revised*. Toronto, Ontario: Multi-Health Systems.
- Hare, R. D. (1998). Psychopaths and their nature: Implications for the mental health and criminal justice systems. In T. Millon, E. Simonsen, M. Birket-Smith, & R. D. Davis (Eds.), *Psychopathy: Antisocial, criminal, and violent behavior* (pp. 188–212). New York: Guilford.
- Hare, R. D. (2003). *Hare Psychopathy Checklist – Revised* (2nd edition). Toronto, Ontario: Multi-Health Systems.
- Hare, R. D., Clarke, D., Grann, M., & Thornton, D. (2000). Psychopathy and the predictive validity of the PCL-R: An international perspective. *Behavioral Sciences & the Law*, 18, 623–645.
- Hart, S. D., & Hare, R. D. (1989). Discriminant validity of the Psychopathy Checklist in a forensic psychiatric population. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 1, 211–218.
- Hart, S. D., & Hare, R. D. (1997). Psychopathy: Assessment and association with criminal conduct. In D. M. Stoff, J. Breiling, & J. D. Maser (Eds.), *Handbook of antisocial behavior* (pp. 22–35). New York: Wiley.
- Heilbrun, K., Hart, S. D., Hare, R. D., Gustafson, D., Nunez, C., & White, A. J. (1998). Inpatient and postdischarge aggression in mentally disordered offenders: The role of psychopathy. *Journal of Interpersonal Violence*, 13, 514–527.
- Hemphill, J. F., & Hart, S. D. (2002). Motivating the unmotivated: Psychopathy, treatment, and change. In M. McMurrin (Ed.), *Motivating offenders to change* (pp. 193–219). Chichester, UK: Wiley.
- Hildebrand, M., & de Ruiter, C. (2004). PCL-R psychopathy and its relation to DSM-IV Axis I and Axis 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., & de Vogel, V. (2004). Psychopathy and sexual deviance in treated rapists: Association with sexual and non-sexual recidivism. *Sexual Abuse: A Journal of Research and Treatment*, 16, 1–24.
- Hildebrand, M., de Ruiter, C., de Vogel, V., & van der Wolf, P. (2002). Reliability and factor structure of the Dutch language version of Hare's Psychopathy Checklist – Revised. *International Journal of Forensic Mental Health*, 1, 139–154.
- Hildebrand, M., de Ruiter, C., & Nijman, H. (2004). PCL-R psychopathy predicts disruptive behavior among male offenders in a Dutch forensic psychiatric hospital. *Journal of Interpersonal Violence*, 19, 13–29.
- Hillege, S., Das, J., & de Ruiter, C. (2010). The Youth Psychopathic traits Inventory: Psychometric properties and its relation to substance use and interpersonal style in a Dutch sample of non-referred adolescents. *Journal of Adolescence*, 33, 83–91.
- Hobson, J., Shine, J., & Roberts, R. (2000). How do psychopaths behave in a prison therapeutic community? *Psychology, Crime, and Law*, 6, 139–154.
- Hughes, G., Hogue, T., Hollin, C., & Champion, H. (1997). First-stage evaluation of a treatment programme for personality disordered offenders. *The Journal of Forensic Psychiatry*, 8, 515–527.
- Jones, M. (1956). The concept of a therapeutic community. *The American Journal of Psychiatry*, 38, 647–650.
- Jones, M. (1968). *Social psychiatry in practice*. Harmondsworth, UK: Penguin.
- Kay, S. R., Wolkenfeld, F., & Murrill, L. M. (1988). Profiles of aggression among psychiatric patients II: Covariates and predictors. *The Journal of Nervous and Mental Disease*, 176, 547–557.
- Kroger, U., van Beek, D., van der Wolf, P., Klein Haneveld, E., van Geest, H., & Geraerts, R. (2011). *Treatment of psychopathy: A Mission Impossible? A treatment program for patients with a high psychopathy score at the Van der Hoeven Hospital*. Utrecht, The Netherlands: Forum Educatief.
- LaForge, R., & Sucek, R. F. (1955). The interpersonal dimension of personality: III. An interpersonal checklist. *Journal of Personality*, 24, 94–112.
- Laws, D. R. (1989). *Relapse prevention with sex offenders*. New York: Guilford.
- Leistico, A. R., Salekin, R. T., DeCoster, J., & Rogers, R. (2008). A large-scale meta-analysis relating the Hare measures of psychopathy to antisocial conduct. *Law and Human Behavior*, 32, 28–45.
- Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. New York: Guilford.
- McNiel, D. E., Eisner, J. P., & Binder, R. L. (2003). The relationship between aggressive attributional style and violence by psychiatric patients. *Journal of Consulting and Clinical Psychology*, 71, 399–403.
- Menzies, R., & Webster, C. D. (1995). Construction and validation of risk assessments in a six year follow-up of forensic patients: A tridimensional analysis. *Journal of Consulting and Clinical Psychology*, 63, 766–778.
- Meyer, G. J. (1997). On the integration of personality assessment method: The Rorschach and MMPI. *Journal of Personality Assessment*, 68, 297–330.
- Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., et al. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist*, 56, 128–165.
- Monahan, J., & Appelbaum, P. S. (2000). Reducing violence risk: Diagnostically based clues from the McArthur Violence Risk Assessment Study. In S. Hodgins (Ed.), *Violence among the mentally ill: Effective treatments and management strategies* (pp. 19–34). Boston, MA: Kluwer.
- Network for Addiction Treatment Novadic (2003). *Interpersoonlijke psychotherapie [Interpersonal psychotherapy]*. [On-line]. Available: <http://www.novadic.nl/interpersoonlijke%20psychotherapie.htm>
- Ogloff, J. R. P., Wong, S., & Greenwood, A. (1990). Treating criminal psychopaths in a therapeutic community program. *Behavioral Sciences & the Law*, 8, 181–190.
- Olver, M. E., & Wong, S. (2009). Therapeutic responses of psychopathic sexual offenders: Treatment attrition, therapeutic change, and long term recidivism. *Journal of Consulting and Clinical Psychology*, 77, 328–336.
- Otto, R. K., & Douglas, K. S. (Eds.). (2010). *Handbook of violence risk assessment*. New York: Routledge.
- Patrick, C. J. (2007). Getting to the heart of psychopathy. In H. Hervé, & J. C. Yuille (Eds.), *The psychopath: Theory, research, and practice* (pp. 207–252). Mahwah, NJ: Erlbaum.
- Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology*, 51, 768–774.
- Pfohl, B., Blum, N., & Zimmerman, M. (1994). *Structured Interview for DSM-IV Personality (SIDP-IV)*. Iowa City: University of Iowa, Department of Psychiatry.
- Pfohl, B., Blum, N., Zimmerman, M., & Stangl, D. (1989). *Structured Interview for DSM-III-R Personality (SIDP-R)*. Iowa City: University of Iowa, Department of Psychiatry.
- Pope, K. S., Butcher, J. N., & Seelen, J. (2006). *The MMPI, MMPI-2 & MMPI-A in court: A practical guide for expert witnesses and attorneys* (3rd ed.). Washington, DC: American Psychological Association.
- Reed, V., Woods, P., & Robinson, D. (2000). *Behavioural Status Index (BEST-Index): A "life skills" assessment for selecting and monitoring therapy in mental health care*. Sheffield, UK: Psychometric Press.
- Rice, M. E., Harris, G. T., & Cormier, C. A. (1992). An evaluation of a maximum security therapeutic community for psychopaths and other mentally disordered offenders. *Law and Human Behavior*, 16, 399–412.
- Rice, M. E., Harris, G. T., Quinsey, V. L., & Cyr, M. (1990). Planning treatment programs in secure psychiatric facilities. In D. N. Weisstub (Ed.), *Law and mental health: International perspectives*, Vol. 5. (pp. 162–230) New York: Pergamon Press.
- Rorschach, H. (1942). *Psychodiagnostics: A diagnostic test based on perception*. Bern: Hans Huber (Original work published 1921).
- Salekin, R. T. (2002). Psychopathy and therapeutic pessimism: Clinical lore or clinical reality? *Clinical Psychology Review*, 22, 79–112.
- Schill, T., & Wang, S. (1990). Correlates of the MMPI-2 anger content scale. *Psychological Reports*, 67, 800–802.
- Seto, M. C., & Barbaree, H. E. (1999). Psychopathy, treatment behavior, and sex offender recidivism. *Journal of Interpersonal Violence*, 14, 1235–1248.
- Skeem, J. (2008, July). *High risk, not hopeless: Recent research on treating individuals with psychopathy*. Paper presented at the annual meeting of the International Association of Forensic Mental Health. Vienna.
- Skeem, J., Monahan, J., & Mulvey, E. (2002). Psychopathy, treatment involvement, and subsequent among civil psychiatric patients. *Law and Human Behavior*, 26, 577–603.
- Sloore, H., Derksen, J. J. L., de Mey, H. R. A., & Hellenbosch, G. (1996). The Flemish/Dutch version of the MMPI-2: Developments and adaptation of the inventory for Belgium and The Netherlands. In J. N. Butcher (Ed.), *International adaptations of the MMPI-2: A handbook of research and clinical application* (pp. 329–349). Minneapolis: University of Minnesota Press.
- Spinella, M. (2007). Normative data and a short form of the Barratt Impulsiveness Scale. *International Journal of Neuroscience*, 117, 359–368.
- Strand, S., Belfrage, H., Fransson, G., & Levander, S. (1999). Clinical and risk management factors in risk prediction of mentally disordered offenders: More important than actuarial data? *Legal and Criminological Psychology*, 4, 67–76.
- van den Brink, W., & de Jong, C. A. J. (1992). *De Nederlandse versie van het Structured Interview for DSM-III-R Disorders of Personality [The Dutch version of the Structured Interview for DSM-III Personality Disorders]*. Eindhoven, The Netherlands: Instituut voor Verslavingszorg Brabant.
- Vertommen, H., Verheul, R., de Ruiter, C., & Hildebrand, M. (2002). *Handleiding bij de herziene versie van Hare's Psychopathie Checklist [Manual of the Dutch version of Hare's Psychopathy Checklist – Revised]*. Lisse, The Netherlands: Swets Test Publishers.
- Vitacco, M. J., Rogers, R., Neumann, C. S., Harrison, K., & Vincent, G. (2005). A comparison of factor models on the PCL-R with mentally disordered offenders: The development of a four-factor model. *Criminal Justice and Behavior*, 32, 526–545.
- Wang, E., & Diamond, P. (1999). Empirically identifying factors related to violence risk in corrections. *Behavioral Sciences & the Law*, 17, 377–389.
- Webster, C. D., Douglas, K. S., Belfrage, H., & Link, B. G. (2000). Capturing change: An approach to managing violence and improving mental health. In S. Hodgins (Ed.), *Violence among the mentally ill: Effective treatment and management strategies* (pp. 119–144). Dordrecht, The Netherlands: Kluwer.
- Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing the risk of violence*. (version 2). Burnaby, BC: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Weiner, I. B. (2003). *Principles of Rorschach interpretation* (2nd ed.). Mahwah, NJ: Erlbaum.
- Weiner, I. B., & Exner, J. E. (1991). Rorschach changes in long-term and short-term psychotherapy. *Journal of Personality Assessment*, 56, 453–465.
- Wong, S. (2000). Psychopathic offenders. In S. Hodgins, & R. Müller-Isberner (Eds.), *Violence, crime and mentally disordered offenders* (pp. 87–112). Chichester, UK: Wiley.
- Young, J. E. (1994). *Cognitive therapy for personality disorders: A schema-focused approach*. Sarasota: Professional Resource Exchange.
- Zamble, E., & Quinsey, V. L. (1997). *The criminal recidivism process*. New York: Cambridge University Press.