Behavior on the ward of personality disordered inpatients and chronically psychotic inpatients during a three-year stay in a Dutch forensic psychiatric hospital

Ruud H.J. Hornsveld¹, Floris W. Kraaimaat², Samantha Bouwmeester³, Machiel A. Polak⁴, & Almar J. Zwets⁵

Number of words (complete manuscript): 8,159
Number of figures: 1
Number of tables: 5

¹ Erasmus University Medical Centre, Rotterdam, NL.
² Radboud University, Nijmegen, NL.
³ Erasmus University, Rotterdam, NL.
⁴ FPC de Kijvelanden, Poortugaal, NL.
⁵ FPC de Kijvelanden, Poortugaal, NL; Erasmus University, Rotterdam, NL.

Correspondence concerning this article should be addressed to Ruud H. J. Hornsveld,
E: r.hornsveld@tiscali.nl
Behavior in Dutch forensic psychiatric inpatients

Abstract

To gain insight into the behavior of a group of personality disordered patients and a group of chronically psychotic patients during their stay in a Dutch forensic psychiatric hospital, data on these patients were collected bi-annually for seven years. Three aspects of the patients' behavior were examined: the prediction of institutional behavior shortly after admission, changes in the patients' behavior on the ward during their stay in hospital, and the prediction of these changes.

In the personality disordered patients, observed irritation/anger and aggressive behavior on the ward turned out to be positively related to psychopathy, the PCL-R lifestyle and antisocial facets, and the neuroticism domain. A positive relationship was also found between aggressive behavior on the ward and trait anger. In the chronically psychotic patients, a positive relationship was found between irritation/anger and the PCL-R interpersonal and lifestyle facet. During a stay of three years, the aggressive behavior of both patient subgroups, which was already low at the start, did not decrease further, but their prosocial behavior increased. In the personality disordered patients, relatively high scores on the antisocial facet of the PCL-R indicated an increase in prosocial behavior, whereas in the chronically psychotic patients no relationship was found between any PCL-R facet and behavior change.

Effect studies on treatment programs for forensic psychiatric inpatients have to contend with the problem of a low base rate of institutional aggression. Therefore, we advise that such studies focus not only on a decrease in negative behaviors but also on an increase in positive behaviors.

Key words: Forensic psychiatry, personality disordered inpatients, chronically psychotic inpatients, behavior on the ward.
Behavior on the ward of personality disordered inpatients and chronically psychotic inpatients
during a three-year stay in a Dutch forensic psychiatric hospital

To gain insight into the behavior of a group of personality disordered patients and a
group of chronically psychotic patients during their stay in a Dutch forensic psychiatric
hospital, data on these patients were collected bi-annually and prospectively for seven years.
Three aspects of the patients’ behavior were examined: the prediction of institutional
behavior shortly after admission, changes in the patients’ behavior on the ward during a
three-year stay in hospital, and the prediction of these changes.

During the past decade, a considerable number of studies devoted to the relationship
between psychopathy and recidivism and to the relationship between psychopathy and
institutional aggression have been published. For instance, Hildebrand, De Ruiter, and
Nijman (2004) investigated the predictive validity of the Psychopathy Checklist-Revised
(PCL-R; Hare, 1991) and its two factors in a group of 92 Dutch forensic psychiatric inpatients
and found that factor 2 (“Chronically unstable and antisocial lifestyle”) was the best predictor
for the total number of institutional incidents, but not for physical violence on the ward. Guy,
Edens, Anthony, and Douglas (2005) performed a meta-analysis of 273 effect sizes to
investigate the associations between the PCL-R and institutional misconduct. Their results
suggested that factor 2 was associated with misconduct more strongly than factor 1 (“Callous
and remorseless use of others”), but for both factors, the associations with physical violence
were smaller than the associations with misconduct. Taking 95 American, Canadian, and
European studies altogether, Leistico, Salekin, DeCoster, and Rogers (2008) performed a
meta-analysis on the relationship between the various editions of the Psychopathy Checklist
(PCL, Hare, 1980) and antisocial conduct, both institutional misbehavior and recidivism after
discharge. The results indicated that higher scores on PCL total, factor 1, and factor 2 were
moderately associated with increased antisocial conduct. However, effect sizes depended on
variables such as country, race, gender, and institutional setting. For instance, the mean
effect sizes of factor 2 were found to be slightly larger in samples of forensic psychiatric
Behavior in Dutch forensic psychiatric inpatients

patients than in samples of detainees. Two years later, Yang, Wong, and Coid (2010) found more or less similar results in comparing the effect sizes of nine risk assessment instruments from 28 original reports, namely that the predictive efficacy of the PCL-R for violence was attributable almost entirely to factor 2. In order to find out whether the combination of factor 1 and factor 2 would result in a better prediction of violence than only factor 2, Kennealy, Skeem, Walters and Camp (2010) conducted a meta-analysis of 32 effect sizes. Again, factor 2 turned out to predict violence better than factor 1, but no interaction was found between the two factors. Although the former two studies supported the evidence for PCL-R factor 2 in being a predictor of future violence, it should be noted that the meta-analyses in these studies overlapped partly with the ones of the Leistico et al. (2008) study and no differentiation was made between forensic psychiatric patients and offenders, and between institutional and community violence.

After Hare’s (2003) introduction of the PCL-R four-facet model, the emphasis shifted from factor 2 to facet 4 in studies on the predictive validity of this instrument (e.g., Walters, Knight, Grann, & Dahle, 2008). In 2010, Walters and Heilbrun published a study on the relationship between facet 4 (“Antisocial”) of the PCL-R and antisocial conduct in a sample of 216 male forensic patients with the PCL (Hare, 1980) and in a sample of 230 psychiatrically evaluated inmates with the PCL-R (Hare, 1991). They found, in accordance with the Walters et al. study (2008), that facet 4 or parcel G (items “early behavioral problems” and “poor behavioral control”) consistently achieved incremental validity to the first three facets (“Interpersonal,” “Affective,” and “Lifestyle”). Although several authors (Hildebrand, De Ruiter, and Nijman, 2004; Leistico et al., 2008; Walters & Heilbrun, 2010) investigated the predictive validity of the PCL-R in (among others) forensic psychiatric patients, no differentiation was made between patients with a major mental disorder on Axis I of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR; American Psychiatric Association, 2000) and those with a primary diagnosis on Axis II in diagnostically mixed samples. Furthermore, Vitacco et al. (2009) suggested that when investigating the utility of the PCL-R and its factors for the prediction of institutional aggression, a differentiation should be made between
reactive and proactive aggression on the ward. In a group of 152 mainly psychotic forensic psychiatric inpatients, they showed that psychopathy (PCL:SV total score) and its Interpersonal and Affective facets were robust predictors for proactive aggression, but not for reactive aggression. However, in explaining the various results of studies several authors have emphasized that the PCL-R has been designed as an instrument for assessing psychopathy and not for assessing the risk of future institutional or community violence. Therefore, they recommend that clinicians and legal decision makers consider risk and protective factors beyond psychopathy when attempting to predict future behaviors (e.g., Edens, Campbell, & Weir, 2007; Leistico, et al., 2008; Singh, Grann, and Fazel, 2011).

Contrary to numerous publications about the relationship between psychopathy and antisocial behavior, the number of publications about the relationship between personality traits as measured by self-report questionnaires and institutional behavior is very limited. Until now, no study has been carried out on the relationship between the Big Five personality domains (Costa & McCrae, 1989) and institutional behavior. The only study about the relationship between anger as a trait and institutional behavior was performed by Cornell, Peterson, and Richards (1999) in a group of 65 incarcerated adolescents. Trait anger, as measured by the State-Trait Anger Expression Inventory (STAXI; Spielberger, 1988) appeared to be significantly correlated with physical and with verbal aggression as rated by the staff.

Studies on the effects of institutional treatment programs for forensic psychiatric patients have yielded divergent results. Recently, McGuire (2013) concluded that cognitive-behavioral treatment programs for violent offenders have positive effects on recidivism, but studies on the behavior change in forensic psychiatric patients during their stay in hospital usually show less convincing results. For instance, Belfrage and Douglas (2002) studied 70 Swedish, mostly psychotic patients with the 20-item Historical/Clinical/Risk Management-20 (Webster, Douglas, Eaves, & Hart, 1997). After an 18-month stay, patients scored significantly lower on the Clinical subscale and the Risk Management subscale. De Jonge, Nijman, and Lammers (2009) had similar results with a study of 158 Dutch, mostly
personality disordered forensic psychiatric patients. After 25 months, patients scored significantly lower on both the Clinical factor and the Future factor of the History/Clinical/Future 30 (Ministry of Security and Justice, 2002), but effect sizes were small. Nijman, De Kruyk, and Van Nieuwenhuizen (2004) studied the course of behavior by means of the Rehabilitation Evaluation Hall and Baker (Baker & Hall, 1988) in a group of 128 Dutch, mainly personality disordered forensic psychiatric inpatients. During hospitalization, social activity improved significantly, but improvements in self-care and verbal skills were not significant. Chakhssi, De Ruiter, and Bernstein (2010) studied a group of 74 Dutch, personality disordered inpatients for a period of 20 months. Using the Behavioral Status Index (BSI; Reed, Woods, & Robinson, 2000), they found that most patients “improved” to a limited extent on the BSI, but the psychopaths (22%) “deteriorated” during the study period, whereas none of the non-psychopaths did. Finally, Hildebrand and De Ruiter (2012) used self-report questionnaires, the Rorschach Inkblot Method (Rohrschach, 1921/1942), staff ratings, and objective measures for treatment compliance for their investigation in a group of 87 Dutch forensic psychiatric inpatients. The group of patients as a whole did not improve on most of the indicators of dynamic risk after 20 months of treatment. A few researchers (Belfrage & Douglas, 2002; Hildebrand & De Ruiter, 2012) suggest that the limited change in the behavior of forensic psychiatric patients during their stay in a hospital might be due to treatment programs, which deserve review or alteration. However, differences in results between the studies about institutional behavior and the studies as summarized by McGuire (2013) might also be explained by differences in outcome measures, namely, reduction of dynamic criminogenic needs versus reduction of recidivism rate.

In the current study, it was hypothesized (1) that psychopathy as assessed by the PCL-R was related to institutional aggression in forensic psychiatric inpatients with a personality disorder, as well as in inpatients with a chronically psychotic disorder, and that this relationship involved specifically facet 4 of the PCL-R; (2) that in each of these patient groups, the hospital stay would result in a decrease of aggressive behavior and in an increase of prosocial behavior; and (3) that these behavior changes were related to
psychopathy in both inpatient groups.

**Method**

**Participants**

In the Netherlands, offenders who have committed a serious violent crime that is punishable with a maximum imprisonment of more than four years (e.g., murder, manslaughter, aggravated assault, or rape) can be detained under hospital order (“TBS order”). This concerns offenders who, based on an extensive psychiatric and/or psychological evaluation at a special assessment center of the Ministry of Security and Justice, are judged to have diminished responsibility for the offense they committed (Van Marle, 2002). TBS involves involuntary admission to a specialized maximum-security forensic psychiatric hospital with obligatory treatment programs that should result in a decrease of recidivism risk to an “acceptable level for society.” The Dutch Ministry of Security and Justice makes a distinction between patients with a “personality disorder” (about 75% of the population) and patients with a “chronically psychotic disorder” (De Beurs & Barendregt, 2008).

Behavior on the ward was assessed prospectively and bi-annually (in the months of January and July) as part of a standard procedure between January 2003 and January 2011 by means of the Observation Scale for Aggressive Behavior (OSAB; Hornsveld, Nijman, Hollin, & Kraaimaat, 2007). During this period, the number of assessed patients gradually increased because of new admissions. Ultimately, data on 253 patients could be collected with one measurement, on 248 patients with two measurements, on 236 patients with three measurements, and so on (Table 1). The mean age of the 253 patients with one measurement (shortly after admission) was 37.46 years ($SD = 10.42$, range: 19-69 years). The 159 patients of the group classified as personality disordered had a mean age of 38.59 years ($SD = 10.67$, range: 19-65 years) at admission, and the 94 classified as chronically psychotic patients had a mean age of 35.54 years ($SD = 9.73$, range: 22-69 years).

Because the average stay of Dutch forensic psychiatric inpatients has increased dramatically in recent years, reaching 9.8 years in 2010 (Nagtegaal, Van der Horst &
Schönberger, 2011), we chose a period of three years for an analysis of behavior on the ward, a compromise between a relatively long period of stay and a sufficient number of patients. In total, 115 patients could be assessed through the OSAB on seven measurement moments from admission onwards. At admission, the 70 personality disordered patients of this group had a mean age of 38.79 years \((SD = 10.75,\text{ range } 19-62\text{ years})\), and the 45 chronically psychotic patients had a mean age of 34.13 years \((SD = 8.85,\text{ range } 22-68\text{ years})\). Some of these 115 patients were admitted in the second half of 2002, were measured for the first time in January 2003, and were measured for the seventh time in January 2006. Other patients were admitted in the first half of 2003, were measured for the first time in July 2003, and were measured for the seventh time in July 2006. The last of these 115 patients were admitted in the second half of 2007, were measured for the first time in January 2008, and were measured for the seventh time in January 2011. Missing scores (e.g., during the fifth measurement) were added by calculating the average score of the two adjacent scores (e.g., during the fourth and sixth measurements). The percentage of missing scores in this part of the data set was 8%.

In the period between January 2003 and January 2008, 204 patients were admitted to the hospital. Data on the remaining 89 patients were incomplete because of several reasons. For 45 patients (group 2), data from less than five measurements could be collected, although they had stayed in the institution for more than three years. Others stayed for less than three years in hospital during the investigated period, because they were discharged (7 patients, group 3), transferred to another institution for re-assessment (20 patients, group 4), or had to continue their care in a long-stay hospital elsewhere (6 patients, group 5). Furthermore, 11 patients (group 6) were admitted only for a very short time, for instance in case of a crisis intervention (Table 2). Groups 2 to 6 did not differ in PCL-R total score from group 1, which included the 115 patients with seven measurements.

**Measures**

The *Observation Scale for Aggressive Behavior* (OSAB; Hornsveld, Nijman, Hollin, & Kraaimaat, 2007) measures behavior on the ward. The scale comprises 40 items spread
over the subscales Irritation/Anger, Anxiety/Gloominess, Aggressive Behavior, Prosocial Behavior, Antecedent, and Sanction. The staff judges the behavior of the inpatients in the preceding week on a four-point scale, with 1 = “no,” 2 = “seldom,” 3 = “occasionally,” and 4 = “frequently.” In this study, we focused on the following subscales: Irritation/Anger (e.g., “Irritated”), Aggressive Behavior (e.g., “Threats toward staff”), and Prosocial Behavior (e.g., “Gives his opinion adequately”). By far, most of these aggressive and prosocial behaviors were exhibited on the ward in relation to staff members. Aggressive behavior towards fellow patients probably took place in the absence of staff members. Aggressive behavior towards fellow patients probably took place in the absence of staff members. In a sample of 220 violent forensic psychiatric inpatients, the internal consistency (Cronbach’s α) was .82, .79, and .93, inter-rater reliability was .79, .81, and .70, and test-retest reliability was .59, .57, and .76, for the subscales Irritation/Anger, Aggressive Behavior, and Prosocial Behavior, respectively (Hornsveld et al., 2007). Convergent validity was supported through significant positive correlations with subscales of the Forensic Inpatient Observation Scale (Timmerman, Vastenburg, & Emmelkamp, 2001).

The Psychopathy Checklist-Revised (PCL-R; Hare, 1991; Dutch version: Vertommen, Verheul, De Ruiter, & Hildebrand, 2002) was employed to measure psychopathy. The checklist consists of 20 items, which have to be rated on a three-point scale, with 0 = “does not apply,” 1 = “applies to some extent,” and 2 = “applies”. Vertommen et al. (2002) found support for the reliability of the Dutch version of the PCL-R in a group of 1,192 inmates. The internal consistency (Cronbach’s α) was .87. Tentative support for convergent validity was found in a subgroup of 98 forensic psychiatric inpatients, as there were modest but meaningful correlations with self-report questionnaires such as the MMPI-2 (Dutch version: Sloore, Derksen, Hellenbosch, & De Mey, 1993). In the present study, we used the total score as well as the four-factor structure as proposed by Hare and Neumann (2006), which implies the following reliable facets: Interpersonal, Affective, Lifestyle, and Antisocial. The PCL-R scores of a sample of 41 patients were assessed independently by two trained psychologists. Intraclass correlation coefficient for the PCL-R total score was .81, and for the four facets .66, .64, .58, and .86 successively. The coefficient of the PCL-R total score was
regarded as excellent, the coefficient of facet 2 as fair, the coefficients of facet 1 and facet 3 as good, and the coefficient of facet 4 as excellent (Cicchetti, 1994).

The NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1992; Dutch version: Hoekstra, Ormel, & De Fruyt, 1996) includes 60 items and measures the Big Five personality domains of neuroticism, extraversion, openness, agreeableness, and conscientiousness. Participants score the items in the NEO-FFI on a five-point Likert scale ranging from “entirely disagree” to “entirely agree.” In the present study, we were only interested in the neuroticism (e.g., “I seldom feel lonely or sad.”) and agreeableness (e.g., “Some people find me selfish and egotistic.”) scales, because these traits are considered relevant in the context of aggression (Hornsveld, Nijman, & Kraaimaat, 2008). In a Dutch sample of 356 non-clinical adults the internal consistency (Cronbach’s α) of the two subscales was .82 and .69, and in a subgroup of 135 adults the test-retest reliability after 6 months was .82 and .75, respectively (Hoekstra et al., 1996).

The Trait Anger subscale of the Spielberger (1980) State-Trait Anger Scale (STAS; Van der Ploeg, Defares, & Spielberger, 1982), which consists of 10 items, was used as a concurrent measure of the general disposition to anger. Participants rate each item about how they generally feel (e.g., “I am quick tempered.”) by using a four-point Likert scale: 1 = “almost never,” 2 = “sometimes,” 3 = “often,” and 4 = “almost always.” In a group of 150 Dutch male university students, Van der Ploeg et al. (1982) found that the internal consistency (Cronbach’s α) of the trait anger scale was .78, and a test-retest reliability of .78 was documented in a subgroup of 70 students. The convergent validity of the trait anger scale also appeared to be satisfactory (Van der Ploeg et al., 1982).

Procedure

The study was approved by the Dutch Review Committee for Patient-linked Research in Arnhem, the Netherlands, and by the Scientific Research and Documentation Center of the Dutch Ministry of Security and Justice.

Group leaders were trained in the scoring of the OSAB, and scored this instrument bi-annually in the months of January and July. Certified clinical psychologists assessed the
PCL-R scores based on file study. Such files comprised detailed information about life history, committed offenses, and elaborated evaluations from psychiatrists and/or psychologists. These reports were drawn up at a special forensic assessment center, in which the offender had to stay for observation by order of the court. The psychologists who scored the PCL-R were not informed about the OSAB-scores.

Patients completed the questionnaires individually under supervision of an experienced research assistant and received a fee of € 7 in return for their participation. Not all patients completed the questionnaires, because participation in the study was on a voluntarily basis.

**Setting**

The current study was conducted at FPC de Kijvelanden in Poortugaal, the Netherlands. The patient-staff ratio was 1 to 1.8. All patients resided on high-security wards for seven to eleven patients, where specifically educated group leaders offered them milieu therapy. During a period of about four months after admission, the patients’ behavior was observed on the ward, psychiatric and psychological evaluations were carried out, and a treatment plan was finally established. Depending on their dynamic criminogenic needs, patients followed cognitive-behavioral treatment programs, focusing on (sexual) violence, addiction, or chronically psychotic disorders. When indicated, they also had to follow individual therapy, creative art therapy (e.g., drama, art, music, or psychomotor therapy), general education, occupational training, or sports. Psychopharmacological therapy was applied to all chronically psychotic patients and to the personality disordered patients for whom it was indicated and who did not refuse medication.

**Statistics**

For the two groups separately, Pearson correlation coefficients were calculated for the relationship between scores on the PCL-R, NEO-FFI, or STAS on the one hand, and the three OSAB subscales shortly after admission on the other hand (two-tailed, \( p < .05 \)). In addition, stepwise multiple regression analyses were performed to explore the relative contribution of the four PCL-R facets to the three OSAB subscales in the two patient groups.
Behavior in Dutch forensic psychiatric inpatients

separately. To study mean linear changes in behavior at the ward as measured by the three OSAB subscales over time (7 measurements, bi-annually) three separate mixed analyses of variance were performed with groups as between-subject factor and the seven measurement moments as within-subject factor. In case of significant relationships, post-hoc tests were performed. To explore the relationship between independent variables at time of admission and changes in behavior at the ward over a period of three years, Pearson correlation coefficients were calculated between the scores on PCL-R, NEO-FFI and STAS at the time of admission and the residual gain scores for the three OSAB scales at measurement 7 (Kerlinger, 1975).

Results

Behavior on the ward shortly after admission

Table 3 shows the mean scores and standard deviations for all measures and 253 patients in each of the two samples.

Personality disordered patients had a significantly higher total score on the PCL-R and its Interpersonal, Affective, and Lifestyle facets than the chronically psychotic patients. No significant difference between subgroups was found on the other measures, with the exception of the Irritation/Anger and Prosocial Behavior subscales of the OSAB, on which the personality disordered patients scored significantly higher than the chronically psychotic patients.

Relationships between OSAB scores on the one hand and PCL-R, NEO-FFI, or STAS scores on the other hand were significant and in the expected direction for the personality disordered patients but not for the chronically psychotic patients. In the personality disordered group, the OSAB subscales Irritation/Anger and Aggressive Behavior correlated significantly positive with the total score on the PCL-R, the Lifestyle facet, the Antisocial facets, and the NEO-FFI domain Neuroticism. The OSAB subscale was also found to correlate significantly positive with the STAS. In the chronically psychotic group, the subscale Irritation/Anger correlated significantly positive with the Interpersonal and Lifestyle facet of the PCL-R. In the latter group, however, no significant relationships were found between the
three OSAB subscales and the two NEO-FFI domains or the STAS (Table 4). Stepwise multiple regression analyses were conducted to evaluate which facets of the PCL-R might predict irritation/anger, aggressive behavior, or prosocial behavior shortly after admission. In the personality disordered group, results indicated that the antisocial facet was significantly related to irritation/anger, $F(1, 157) = 14.42, p < .01$. The multiple correlation coefficient was .29, indicating that approximately 8.4% of the variance of irritation could be accounted for by the antisocial facet scores. The other facets of the PCL-R were not entered into the equation (all $p$'s > .33). Aggressive behavior was only significantly related to the lifestyle facet in the group of personality disordered patients, $F(1, 157) = 12.71, p < .01$. The multiple correlation coefficient was .27, indicating that approximately 7.5% of the variance of aggressive behavior could be accounted for by the lifestyle facet scores. The other facets of the PCL-R were not entered into the equation (all $p$'s > .62). None of the facets of the PCL-R was significantly related to prosocial behavior in the group of personality disordered patients. In the group of chronically psychotic patients, results indicated that the lifestyle facet of the PCL-R was significantly related to irritation, $F(1, 92) = 5.89, p < .02$. The multiple correlation coefficient was .25, indicating that approximately 6.0% of the variance of irritation could be accounted for by the lifestyle facet score. The other facets were not entered into the equation (all $p$'s > .24). For the group of psychotic patients, none of the facets of the PCL-R was significantly related to aggressive behavior or prosocial behavior (Table 5).

**Change in behavior on the ward over a period of three years**

The change of behavior was studied in the 115 patients (70 personality disordered and 45 chronically psychotic patients) with seven measurements. The 45 patients who also stayed three years or more in the hospital but from whom less than five data measurements could be collected (Table 2) did not significantly differ from the 115 patients with seven measurements in PCL-R total score. Therefore, we assume that the studied group of 115 patients was representative of all patients who stayed three years or longer in the hospital.

For the scores on the Irritation/Anger subscale of the OSAB, both the assumptions of homogeneity of variance matrices ($(\text{Box's } M = 33.58, F(28, 30926.168) = 1.25, p = .17)$) and
sphericity (Greenhouse Geisser $\varepsilon = .87$) were met. The effect of the measurement within-subject factor Irritation/Anger was significant ($\text{(Wilks' Lambda } F(6, 108) = 5.107, p < .001)$) and the effect size was large (partial $\eta^2 = .221$). Both the between-subject factor group ($\text{(}F(1, 113) = 4.682, p = .033, \text{ partial } \eta^2 = .040\text{)}$) and the interaction effect group x irritation/anger measurement ($\text{(Wilks' Lambda } F(6, 108) = 2.031, p = .068, \text{ partial } \eta^2 = .101\text{)}$) were not significant. A polynomial contrast analysis showed a significant quadratic trend effect, $F(1, 113) = 16.402, p < .001, \text{ partial } \eta^2 = .127$. The observed averages of irritation/anger on each time in Figure 1 per group show that patients on average became more irritated/angry during the first measurements, then their irritation dropped and slightly increased on the last measurement.

Regarding the scores on the OSAB subscale Aggressive Behavior, both the assumptions of homogeneity of variance matrices ($\text{(Box's M = 38.04, } F(28, 30926.168) = 1.263, p = .16\text{)}$) and sphericity (Greenhouse Geisser $\varepsilon = .81$) were also met. None of the effects were significant ($\text{(within-subject factor: Wilks' Lambda } F(6, 108) = 1.056, p = .394, \text{ partial } \eta^2 = .055; \text{ between-subject factor: } F(1, 113) = 0.25, p = .620, \text{ partial } \eta^2 = .002; \text{ interaction aggressive behavior x group: Wilks' Lambda } F(6,108) = .988, p = .437, \text{ partial } \eta^2 = .052\text{)}$). Figure 1 shows the observed average aggressive behavior level per group.

For the scores on the OSAB subscale Prosocial Behavior, the assumption of homogeneity of variance matrices was violated, Box's M = 50.63, $F(28, 30926.168) = 1.680, p = .014$. Levene's test of equality of variances shows that the variances on measurement 2, 3, 5, and 7 were unequal, with more variance in the chronically psychotic group than in the personality disordered group. Because the psychotic group is much smaller than the personality disordered group, the $F$-test of the between-subject factor group and the interaction groups x prosocial behavior is liberal (Stevens, 2007, p. 58). We therefore decided to use an alpha level of .01. The sphericity assumption was just about met (Greenhouse Geisser $\varepsilon = .76$). The effect of the within-subject factor Irritation measurement was significant ($\text{(Wilks' Lambda } F(6, 108) = 4.352, p = .001\text{)}$), but the effect size was small
Behavior in Dutch forensic psychiatric inpatients

(partial $\eta^2 = .195$). The between-subject factor was also significant, $F(1, 113) = 20.80, p < .001$, partial $\eta^2 = .155$. On average, the personality disordered group scored higher on prosocial behavior than the chronically psychotic group. The interaction effect group x prosocial behavior measurement ($\text{(Wilks' Lambda } F(6, 108) = 0.864, p = .52, \text{partial } \eta^2 = .046)$) was not significant. A polynomial contrast analysis showed a significant quadratic trend effect, $F(1, 113) = 8.69, p = .004$, partial $\eta^2 = .710$. The observed averages of prosocial behavior at each time point in Figure 1 show that patients on average showed more prosocial behavior during the first measurements, that their prosocial behavior then dropped and increased slightly again on the last measurement.

Because of the decreasing number of patients with an increasing number of measurement moments (see Table 1) and in order to corroborate our findings with the mixed analyses of variance, a comparison was performed on the OSAB subscales Irritation/Anger, Aggressive Behavior, and Prosocial Behavior with paired $t$-tests between measurement 2 and measurement 1 ($n = 248$), between measurement 3 and measurement 1 ($n = 236$), and so on to measurement 7 and measurement 1 ($n = 115$, see also Table 1). Similar results were found with these $t$-tests as with the mixed analyses of variance.

**Predictors of behavior change after three years**

During measurement 7, no relationships were found between PCL-R and OSAB scores in the personality disordered group, but in the chronically psychotic group both irritation/anger ($r = .38, p < .05$) and prosocial behavior ($r = .33, p < .05$) turned out to be related with the Lifestyle facet of the PCL-R.

When we correlated residual change scores (measurement 7 versus measurement 1) on the three OSAB subscales with PCL-R, NEO-FFI, and STAS, we found that higher scores on the PCL-R facet Antisocial in the personality disordered patients indicated a larger increase in prosocial behavior ($r = .29, p < .05$). In the chronically psychotic patients, no relationships were found between PCL-R facets and behavior change.
Behavior on the wards of a forensic psychiatric institution was studied in a group of personality disordered patients and a group of chronically psychotic patients. In accordance with the findings of Walter and Heilbrun (2010), in the personality disordered patients institutional aggression shortly after admission turned out to be related to the antisocial facet of psychopathy. However, in the chronically psychotic group a relationship was found only between irritation/anger and the psychopathy lifestyle facet. After three years, no relationships between institutional aggression and any PCL-R facet were found in the personality disordered group. In the chronically psychotic group, however, the psychopathy lifestyle facet again turned out to be related to irritation/anger and this point in time also to prosocial behavior. That no relationship between aggression on the ward and any PCL-R facet was found in the personality disordered patients after a stay of three years in hospital supports the findings of Heilbrun et al. (1998). In a group of 218 mentally disordered offenders, these researchers found significant correlations between aggressive incidents and PCL-R scores during the first two months of their stay in hospital, but these associations were no longer significant during the last two months of hospitalization. Regarding the different findings in the chronically psychotic group, it should be noted that both patient groups stayed in a similar controlled and structured environment. Probably, the personality disordered patients in this study are more similar to North American prison samples and the chronically psychotic patients more similar to North American forensic psychiatric samples with especially schizophrenic or other psychotically disordered patients (Hildebrand & De Ruiter, 2012).

In both groups, irritation/anger and aggression on the ward were already low shortly after admission and did not decrease further during their stay of three years in hospital. However, prosocial behavior increased significantly in the two groups during this period. In the personality disordered patients, the psychopathy antisocial facet was found to be related to a change in prosocial behavior, whereas none of the PCL-R facets was related to irritation/anger, aggressive behavior, or prosocial behavior in the chronically psychotic group.
Our study had a number of limitations. First, PCL-R scores were based on file study alone and most scores were single rated. In a sample of 41 inpatients, the intraclass correlation coefficient of the PCL-R total score was excellent, but the coefficients of the four facets varied from fair to excellent. Second, in 8% of the cases that were studied on behavior change during a three-year period, we had to interpolate data for the observation scale because of missing values. Third, some patients refused to complete the self-report questionnaires, and the patients who responded might have done this in a more or less socially desirable way. Consequently, the number of assessed patients was rather low for some measures. Fourth, the validity of the NEO-FFI and STAS has not been studied in Dutch forensic psychiatric patients until now. Finally, with the observation scale no difference could be made between reactive and proactive aggression. However, because most aggressive behavior was exhibited in interactions with the staff, we suppose that this aggressive behavior was mostly reactive of sorts. This would explain why no relationships were found between the psychopathy interpersonal and affective facets as in the Laurell, Belfrage, and Hellström (2010) study.

In spite of these limitations, several provisional conclusions can be drawn. For example, the structured and controlled environment of the forensic psychiatric hospital probably had an attenuating effect on the aggressive behavior of most patients. This effect is in line with earlier studies, in which inpatients turned out to score lower than forensic psychiatric outpatients on multiple-choice self-report questionnaires measuring anger and aggression (Hornsveld, Muris, & Kraaimaat, 2011; Hornsveld, Muris, Kraaimaat, & Meesters, 2009). Vitacco et al. (2009) explained the various results of the studies on institutional behavior through the insufficient uniformity in definitions of institutional aggression and the lack of differentiation between reactive and proactive aggression. They also noticed that base rates of institutional aggression are generally low, due to medication and the presence of staff.

Our findings and those of others may illustrate the problems by demonstrating the effects of treatment programs for inpatients when the outcome is based only on negative
behavior, such as hostility, anger, or aggression. It is therefore recommended to study positive, socially desirable behavior additionally. A recent study by De Vries Robbé, De Vogel, and Douglas (2013) indicates that protective factors, which mostly refer to positive behavior, may contribute to a more valid assessment of recidivism risk.

In our opinion, the present study may indicate that for a considerable group of patients a three- to four-year stay in a Dutch forensic psychiatric institution should suffice for assessment of risk and needs, establishing a treatment plan, and executing this plan through treatment programs. Therefore, with the average stay of 9.8 years for Dutch forensic psychiatric inpatients in mind, it is probably worthwhile to investigate whether this period can be shortened by transferring these patients to a less structured environment, in order to safely practice their newly acquired behavior. Such an environment would also allow professionals to monitor the patients' behavior, assess treatment results, and intervene in case of an unexpected relapse. Further research may indicate if a policy focusing on a safe but more gradual return of those inpatients in society is not only cheaper, but also increases treatment efficacy.
References


Swets Test Services.


Table 1

Number of patients with number of measurement moments, admitted between January 2003 and January 2008.

<table>
<thead>
<tr>
<th>Measurement moments</th>
<th>Total group of patients</th>
<th>Personality disordered patients</th>
<th>Chronically psychotic patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Age</td>
<td>n</td>
</tr>
<tr>
<td>1</td>
<td>253</td>
<td>37.46 (10.42)</td>
<td>159</td>
</tr>
<tr>
<td>2</td>
<td>248</td>
<td>37.40 (10.44)</td>
<td>157</td>
</tr>
<tr>
<td>3</td>
<td>236</td>
<td>37.17 (10.41)</td>
<td>148</td>
</tr>
<tr>
<td>4</td>
<td>213</td>
<td>37.20 (10.34)</td>
<td>134</td>
</tr>
<tr>
<td>5</td>
<td>178</td>
<td>36.97 (10.46)</td>
<td>108</td>
</tr>
<tr>
<td>6</td>
<td>146</td>
<td>37.03 (10.17)</td>
<td>86</td>
</tr>
<tr>
<td>7</td>
<td>115</td>
<td>36.97 (10.27)</td>
<td>70</td>
</tr>
<tr>
<td>8</td>
<td>84</td>
<td>36.68 (9.57)</td>
<td>54</td>
</tr>
<tr>
<td>9</td>
<td>70</td>
<td>37.03 (9.66)</td>
<td>47</td>
</tr>
<tr>
<td>10</td>
<td>48</td>
<td>35.75 (8.91)</td>
<td>30</td>
</tr>
<tr>
<td>11</td>
<td>24</td>
<td>36.50 (8.40)</td>
<td>16</td>
</tr>
</tbody>
</table>
Table 2

**Mean scores and standard deviations for age, PCL-R, and OSAB in all patients who were admitted in the hospital from January 2003 to January 2008.**

<table>
<thead>
<tr>
<th>Number</th>
<th>Group</th>
<th>Number of patients</th>
<th>Percentage</th>
<th>Age</th>
<th>PCL-R Total</th>
<th>OSAB Irritation/anger</th>
<th>Aggressive behavior</th>
<th>Prosocial behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Three year stay in hospital with seven measurements</td>
<td>115</td>
<td>56.4</td>
<td>36.97 (10.27)</td>
<td>20.49 (7.97)</td>
<td>10.19 (3.62)</td>
<td>14.97 (5.66)</td>
<td>17.43 (8.60)</td>
</tr>
<tr>
<td>2</td>
<td>Three year stay in hospital with less than five measurements</td>
<td>45</td>
<td>22.1</td>
<td>38.62 (11.95)</td>
<td>18.00 (7.87)</td>
<td>11.32 (3.76)</td>
<td>15.92 (4.60)</td>
<td>29.11 (7.73)</td>
</tr>
<tr>
<td>3</td>
<td>End of TBS or on leave</td>
<td>7</td>
<td>3.4</td>
<td>41.29 (11.94)</td>
<td>17.00 (8.25)</td>
<td>10.14 (2.48)</td>
<td>14.14 (2.48)</td>
<td>35.29 (4.50)</td>
</tr>
<tr>
<td>4</td>
<td>Transfer or reselection</td>
<td>20</td>
<td>9.8</td>
<td>36.95 (8.89)</td>
<td>22.63 (7.86)</td>
<td>11.63 (3.67)</td>
<td>16.16 (5.23)</td>
<td>27.89 (7.64)</td>
</tr>
<tr>
<td>5</td>
<td>Long-stay hospital</td>
<td>6</td>
<td>2.9</td>
<td>55.00 (8.46)</td>
<td>23.40 (7.57)</td>
<td>13.20 (2.49)</td>
<td>18.40 (2.61)</td>
<td>30.80 (5.98)</td>
</tr>
<tr>
<td>6</td>
<td>Very short stay</td>
<td>11</td>
<td>5.4</td>
<td>35.89 (6.31)</td>
<td>25.67 (9.35)</td>
<td>11.22 (2.59)</td>
<td>15.67 (3.20)</td>
<td>27.44 (6.00)</td>
</tr>
</tbody>
</table>

*Note: PCL- R = Psychopathy Checklist-Revised; OSAB = Observation Scale for Aggressive Behavior.*
### Table 3

*Mean scores and standard deviations for PCL-R, NEO-FFI, STAS, and OSAB in personality disordered (n = 159) and chronically psychotic patients (n = 94).*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Factor or subscale</th>
<th>Personality disordered</th>
<th>Chronically psychotic</th>
<th>Differences between subgroups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M (SD)</td>
<td>N</td>
<td>M (SD)</td>
</tr>
<tr>
<td>Age</td>
<td>159</td>
<td>38.59 (10.67)</td>
<td>94</td>
<td>35.54 (9.73)</td>
</tr>
<tr>
<td>PCL-R</td>
<td>Psychopathy</td>
<td>159</td>
<td>22.25 (8.06)</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Interpersonal</td>
<td>159</td>
<td>3.57 (2.47)</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Affective</td>
<td>159</td>
<td>6.13 (1.72)</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Lifestyle</td>
<td>159</td>
<td>5.94 (2.76)</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Antisocial</td>
<td>159</td>
<td>5.11 (2.81)</td>
<td>94</td>
</tr>
<tr>
<td>NEO-FFI</td>
<td>Neuroticism</td>
<td>97</td>
<td>32.24 (8.49)</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>Agreeableness</td>
<td>97</td>
<td>41.59 (5.23)</td>
<td>48</td>
</tr>
<tr>
<td>STAS</td>
<td>Trait anger</td>
<td>92</td>
<td>17.91 (6.64)</td>
<td>47</td>
</tr>
<tr>
<td>OSAB</td>
<td>Irritation/anger</td>
<td>159</td>
<td>11.08 (3.42)</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Aggressive behavior</td>
<td>159</td>
<td>15.61 (5.05)</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Prosocial behavior</td>
<td>159</td>
<td>29.95 (7.80)</td>
<td>94</td>
</tr>
</tbody>
</table>

*Note: PCL-R = Psychopathy Checklist-Revised; NEO-FFI = Five Factor Inventory; STAS = State-Trait Anger Scale; OSAB = Observation Scale for Aggressive Behavior.*
Table 4

Correlations (Pearson’s r) between scores on PCL-R, NEO-FFI, STAS, and OSAB for personality disordered (n = 159) and chronically psychotic patients (n = 94), assessed shortly after admission (measurement 1).

<table>
<thead>
<tr>
<th>Measure</th>
<th>Factors or subscales</th>
<th>Personality disordered patients</th>
<th>Chronically psychotic patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>Irritation/ Anger</td>
</tr>
<tr>
<td>PCL-R</td>
<td>Psychopathy</td>
<td>159</td>
<td>.24**</td>
</tr>
<tr>
<td></td>
<td>Interpersonal</td>
<td>159</td>
<td>.10</td>
</tr>
<tr>
<td></td>
<td>Affective</td>
<td>159</td>
<td>.18*</td>
</tr>
<tr>
<td></td>
<td>Lifestyle</td>
<td>159</td>
<td>.21**</td>
</tr>
<tr>
<td></td>
<td>Antisocial</td>
<td>159</td>
<td>.29**</td>
</tr>
<tr>
<td>NEO-FFI</td>
<td>Neuroticism</td>
<td>97</td>
<td>.20*</td>
</tr>
<tr>
<td></td>
<td>Agreeableness</td>
<td>97</td>
<td>-.10</td>
</tr>
<tr>
<td>STAS</td>
<td>Trait anger</td>
<td>92</td>
<td>.14</td>
</tr>
</tbody>
</table>

*Note: PCL-R = Psychopathy Checklist-Revised; NEO-FFI = Five Factor Inventory; STAS = State-Trait Anger Scale; OSAB = Observation Scale for Aggressive Behavior. * p < .05, ** p < .01 (two-tailed).
Table 5

*Stepwise multiple regression analyses between facets of PCL-R, and OSAB subscales for personality disordered (n = 159) and chronically psychotic patients (n = 94), assessed shortly after admission (measurement 1).*

<table>
<thead>
<tr>
<th>OSAB Subscale</th>
<th>Personality disordered patients</th>
<th>Chronically psychotic patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PCL-R</td>
<td>B</td>
</tr>
<tr>
<td>Irritation/anger</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facet 4</td>
<td>0.354</td>
<td>0.093</td>
</tr>
<tr>
<td>Aggressive behavior</td>
<td>Facet 4</td>
<td>0.493</td>
</tr>
<tr>
<td>Prosocial behavior</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*Note:* PCL-R = Psychopathy Checklist-Revised; OSAB = Observation Scale for Aggressive Behavior.
Figure 1

Course of behavior during the first three years of stay for 70 personality disordered and 45 chronically psychotic patients, measured with the OSAB subscales Irritation/anger, Aggressive behavior, and Prosocial behavior.
Note: Irritation/anger scores run from 9 to 36, Aggressive behavior scores from 10 to 40, and Prosocial scores from 12 to 48.