Screening for Psychopathy: Validation of the Psychopathic Personality Inventory–Short Form with Reference Scores

Some of the authors of this publication are also working on these related projects:

Help-seeking behavior in victims of domestic violence View project

Interpersonal behavior of psychopathic individuals View project

Screening for Psychopathy: Validation of the Psychopathic Personality Inventory-Short Form with Reference Scores

**Franca Tonnaer · Maaike Cima · Klaas Sijtsma ·
Kasia Uzieblo · Scott O. Lilienfeld**

Abstract The current study reports validation results for the Psychopathic Personality Inventory (PPI) and its subscales, and for a newly developed PPI-Short Form (PPI-SF) in forensic and non-forensic populations. We also provide criterion reference scores for the PPI and the PPI-SF. In Study 1, we used PPI data from 1,065 participants and supplementary PCL-R data from a subsample of 91 forensic offenders. Mokken scale analysis was used to construct the PPI-SF. In Study 2, PPI-SF and PCL-R data were collected from 60 participants. The study yielded promising but preliminary support for the construct validity of the PPI and the PPI-SF. The PPI-SF is of interest for risk assessment because of its (a) strong relationship with the PCL-R total score and (b) subscales known for their predictive value for violence and criminal recidivism.

F. Tonnaer (✉)
Department of Research, Forensic Psychiatric Centre De Rooyse Wissel, P.O. Box 433, 5800 AK Venray, The Netherlands
e-mail: FTonnaer@deRooyseWissel.nl

F. Tonnaer
Department of Clinical Psychological Science,
Maastricht University, Maastricht University, P.O. Box 433,
5800 AK Venray, The Netherlands

M. Cima
Department of Developmental, Clinical and Crosscultural Psychology, Tilburg University, Tilburg, The Netherlands

M. Cima
Department of Research, Forensic Psychiatric Centre De Rooyse Wissel, Venray, The Netherlands

K. Sijtsma
Department of Methodology and Statistics, Tilburg University, Tilburg, The Netherlands

K. Uzieblo
Department of Applied Psychology, College University Lessius, Antwerp, Belgium

S. O. Lilienfeld
Department of Psychology and Interdisciplinary Sciences,
Emory University, Atlanta, USA

Psychopathy is a condition marked by a pattern of superficial charm, callousness, dishonesty, egocentricity, and unmotivated antisocial behavior (Hare 1991). The Psychopathy Checklist-Revised ([PCL-R] Hare 2003) is the most commonly used measure to assess psychopathy. Professionals score the PCL-R based on an interview and collateral file records. The instrument measures personality traits and behaviors, including glibness, impulsivity, aggressive actions, and lack of empathy. The PCL-R is a reliable and valid instrument for assessing psychopathy (Morrissey et al. 2005) but it has some pragmatic limitations. First, the use of the PCL-R requires extensive training and practice. Second, the PCL-R is time-consuming, which is a potential limitation in many forensic settings, in which the time allotted to obtain a diagnosis and conduct a risk assessment is typically restricted. Third, collateral files often are unavailable outside of prison or other institutions, rendering PCL-R assessment impractical (Lilienfeld and Andrews 1996).

In an effort to overcome these pragmatic limitations, Lilienfeld and Andrews (1996) developed the Psychopathic Personality Inventory (PPI) to assess the core features of psychopathy in non-institutionalized settings. The PPI is a self-report instrument developed and validated in undergraduate samples, and its total score is moderately to highly related to PCL-R total scores in prison samples (Poythress et al. 1998). As the PPI was developed primarily as a research tool, it lacks criterion reference tables for the interpretation

of individual scores. Later, Lilienfeld and Widows (2005) developed a revised version of the PPI (the PPI-R) to eliminate psychometrically suboptimal items, lower its reading level, and increase the measure's cross-cultural applicability.

The goals of this study were to contribute to the validation of the PPI and to construct and validate a new PPI-Short Form (PPI-SF) with reference scores. A short form of the PPI has been developed previously (e.g., Lilienfeld and Hess 2001; Vaughn, Howard and DeLisi 2008; Vaughn, Newhill et al. 2008), but the potential disadvantage of this short form is that it was developed by selecting subsets of 7 items each that loaded the most highly on each of 8 dimensions identified in a factor analysis of the full PPI. This strategy is defensible, but it may result in scales that are overly narrow in content, because items with the highest factor loadings will typically be those that display the highest inter-item correlations (stemming from high levels of substantive overlap among items; Smith et al. 2000). Another disadvantage is that this approach may exclude items that are rarely endorsed but highly discriminating at either the low or high end of the psychopathy continuum. One advantage of Mokken scale analysis (MSA), which was used in the current study, is that it may identify items with greater breadth better reflecting the 'core' of a construct, in this case psychopathy. In this respect, this methodological approach may produce a measure superior in some ways to not only the PPI-SF but also the full PPI and PPI-R, as it may target items with better discrimination power at both the low and high ends of the psychopathy continuum.

The current study aims at developing a well-validated PPI-SF with criterion reference scores. To find out whether the PPI discriminates among psychopathic, non-psychopathic, and healthy comparison participants, in the first study we investigated the psychometric qualities of the PPI. The second study focused on the psychometric properties of the new PPI short form (PPI-SF).

Study 1

The first study had three aims. First, we examined the reliability and the validity of the PPI. Second, we developed criterion reference scores. Third, we designed the PPI-SF.

Method

Participants

Over a period of three years, we collected PPI scores from 1,065 male participants (588 forensic patients and 477 healthy controls). The control sample included a student sample ($n=311$) and a community sample ($n=166$). The

ethical commission of Maastricht University approved the research protocol. Participants cooperated on a voluntary basis. They received written and oral instructions emphasizing that participation was not related to prospects for parole and that they were free to withdraw from the study at any time. Participants ranged in age from 13 to 72 years ($M=31.3$, $SD=12.2$), with a mean age of 34.0 ($SD=11.6$) for the forensic patients, and 28.1 ($SD=12.1$) for the healthy controls. Forensic participants were older than the controls, $t(1053)=8.07$, $p<.001$, $d=.50$. Educational information was available for most participants (72 %). Of this majority, 11 % attended elementary school and 38 % secondary school; 51 % had a bachelor-level education. For forensic individuals, the corresponding percentages of these three levels of education were 27 %, 70 %, and 3 % respectively, among healthy controls, they were 0 %, 15 %, 84 %, respectively. These findings indicated higher education levels for control participants, $t(762)=34.5$, $p<.001$, $d=2.5$. For 57 % of the forensic participants, offenses were known; 16 % were convicted for (attempted) manslaughter or murder, 15 % for bodily harm, 14 % for property crimes, 13 % for property crimes with violence, 11 % for drug and weapon related crimes, 7 % for sexual offenses with minors, 6 % for sexual offences, 3 % for deprivation of freedom, and 2 % for arson; 12 % were convicted for other crimes. From the current sample, PPI scores of 167 male graduates and 165 prisoners were reported previously (Uzieblo et al. 2007; Verschuere et al. 2007). Supplementary PCL-R data were available for a subsample ($n=91$) of forensic offenders, which were retrieved for diagnostic purposes. Two trained forensic professionals scored and discussed all PCL-R interviews, resulting in a consensus score based on discussion.

Measures

Psychopathic Personality Inventory ([PPI] Lilienfeld and Andrews 1996) The PPI contains 187 self-report items, each of which is answered on a 4-point Likert scale (1=*false* and 4=*true*). The PPI assesses a variety of psychopathic personality characteristics that are grouped into eight factor-analytically developed content subscales: (1) Machiavellian Egocentricity (30 items) assesses ruthlessness and a willingness to manipulate others, (2) Social Potency (24 items) assesses charm and interpersonal dominance, (3) Coldheartedness (21 items) assesses callousness and an absence of guilt, (4) Carefree Nonplanfulness (20 items) assesses a failure to plan behavior and inhibit maladaptive impulses, (5) Fearlessness (19 items) assesses a propensity for risk taking behavior, (6) Blame Externalization (18 items) assesses externalizing misbehavior, (7) Impulsive Nonconformity (17 items) assesses a lack of concern about social traditions, and (8) Stress Immunity (11 items) assesses the absence of emotional reactions to potentially anxiety-provoking

events. Internal consistency of the PPI content scales is adequate to high (alphas range from 0.70 to 0.93), and construct validity is supported by moderate to high associations with other measures of psychopathy and low associations with nonspecific measures of generalized maladjustment and social desirability (Lilienfeld and Andrews 1996; Lilienfeld and Widows 2005).

Psychopathy Checklist-Revised ([PCL-R] Hare 1991, 2003) The PCL-R is a 20-item measure of psychopathy based on an interview and collateral file records. Each item is evaluated on a 3-point scale (0=*does not apply* and 2= *does definitely apply*). Hare (1991) proposed a two-factor structure consisting of an Affective-Interpersonal factor and a Behavioral Lifestyle factor. However, the most recent version of the PCL-R manual presented a four-factor model (Hare 2003). The first (Interpersonal) factor comprises symptoms such as glibness and grandiose self-worth. The second (Affective) factor comprises symptoms such as lack of remorse or guilt, and irresponsibility. The third (Behavioral Lifestyle) factor assesses features such as impulsivity and lack of long-term goals. The fourth (Antisocial) factor comprises poor behavioral control, juvenile delinquency, and multifarious criminality. The internal consistency of the PCL-R is adequate (alpha for the total score ranged from .83 to .87, Hildebrand 2008). An extensive body of self-report, interview, and psychophysiological research supports the PCL-R's construct validity (Hare 2003).

Statistical Analysis

The distributions of the PPI total and the subscale scores were essentially normal[1] (see Table 1). Only 18 participants had more than 3 missing item scores. Two-way missing-value imputation (Van Ginkel and Van der Ark 2005) was used to estimate the missing item scores. The internal consistency of the PPI was estimated using coefficient alpha. Group differences between forensic patients and healthy controls were investigated using the independent-samples *t*-test. Convergent validity was investigated using Pearson product–moment correlations. Moreover, PPI criterion reference scores were collected in the total sample ($N=1065$). Mokken scale analysis (MSA) using the program MSP (Molenaar and Sijtsma 2000) investigated the scalability for individual items and entire scale, estimated item popularity, and item discrimination.

MSA is a frequently used scaling procedure but may require some introduction (see Emons et al. 2012). MSA is based on a nonparametric item response theory (IRT) model (Sijtsma and Molenaar 2002). Compared with factor analysis

(FA) and principal components analysis (PCA), MSA possesses several advantages. One advantage is that MSA is well suited for discrete item scores, such as rating scale scores, whereas FA and PCA assume item scores to be continuous. MSA uses the scalability coefficient H for dimensionality analysis and item selection. The researcher chooses a minimum value c for H and an automatic item selection procedure selects only those items in clusters that produce a scale for which $H \geq c$. We investigated the item dimensionality by selecting different values for c starting with $c=.30$ and raising c by .05 in each next selection round until $c=.50$. Ultimately, studying the pattern of cluster outcomes across the different c values indicates the true dimensionality of items (Sijtsma and Molenaar 2002, chap. 6). Items $H<.3$ are unscalable (Sijtsma and Molenaar 2002, p. 60), $.3 \leq H < .4$ indicates a weak scale, $.4 \leq H < .5$ a medium scale, and $H \geq .5$ a strong scale (Mokken, 1971, p. 185). We used mean item scores as item locations, indicating the popularity of the specific behavioral aspect with respect to the personality scale. Item coefficient H_j indicates the extent to which item j is related to the other items in the scale, with higher H_j values indicating better discriminating power and greater contributions to a reliable person ordering.

Results

Reliability of the PPI

In the total sample, total PPI score coefficient alpha was .92. For the PPI subscales, alphas ranged from .70 (Stress Immunity) to .87 (Carefree Nonplanfulness). For the total PPI score, coefficient alpha was .92 among forensic participants and .91 among controls. For the forensic participants, alpha for the PPI subscales ranged from .67 (Stress Immunity) to .88 (Carefree Nonplanfulness). For the controls, alpha for the PPI subscales ranged from .74 (Stress Immunity) to .86 (Carefree Nonplanfulness and Blame Externalization).

Validity of the PPI

Group Differences Independent-samples *t*-tests and effect sizes (using Cohen's *d*) were computed to examine whether forensic participants and controls differed on the PPI subscales. Table 1 shows that the forensic group had a significantly higher mean total score than the controls. The forensic group scored significantly higher on the PPI subscales of Fearlessness, Blame Externalization and Stress Immunity. Unexpectedly, controls scored significantly higher on Impulsive Nonconformity. The groups did not differ significantly on Machiavellian Egocentricity, Social Potency, Coldheartedness, and Carefree Nonplanfulness.

---

[1] Because of acceptable skewness and kurtosis values, no extreme scorers were excluded.

Table 1 Group Differences for PPI Total Scale and Subscales Between Forensic and Control Groups; Skewness and Kurtosis, Independent-Samples t-Tests, and Effect Size (d)

	Forensic		Controls		Skewness	Kurtosis	t	d
	M	SD	M	SD				
Total score	388 .86	(48 .4)	378 .86	(42 .4)	0.06	0.15	3.59[***]	0.22
Machiavellian Egocentricity	66 .54	(14 .2)	65 .96	(11 .5)	0.14	0.01	0.75	0.05
Social Potency	61 .88	(10 .9)	61 .05	(8 .6)	0.15	0.59	1.39	0.09
Coldheartedness	49 .73	(12 .7)	50 .37	(10 .1)	0.17	0.65	0.92	0.06
Carefree Nonplanfulness	44 .36	(12 .4)	45 .69	(10 .6)	0.08	0.90	1.89	0.12
Fearlessness	48 .24	(10 .3)	46 .43	(9 .2)	0.03	0.20	3.04[**]	0.19
Blame Externalization	45 .39	(9 .9)	35 .69	(8 .8)	0.21	0.57	16.94[***]	1.04
Impulsive Nonconformity	37 .46	(8 .1)	39 .13	(8 .3)	0.10	0.40	3.31[***]	0.20
Stress Immunity	29 .19	(6 .0)	28 .38	(5 .5)	0.05	0.45	2.30[*]	0.14
Deviant Responding	23 .44	(3 .0)	24 .33	(2 .2)	0.25	1.76	5.53[***]	0.34
Unlikely Virtues	37 .02	(4 .7)	37 .75	(3 .9)	0.12	0.09	2.58[**]	0.16
VRIN	1 .47	(8 .0)	2 .65	(6 .7)	0.04	0.08	2.62[**]	0.16

[*] $p < .05$, [**] $p < .01$, [***] $p < .001$

Convergent Validity In the forensic subsample, we found significant correlations between the PPI total score and PCL-R total score, $r = .40$, $.21 < r < .56$, the PCL-R Lifestyle factor, $r = .48$, $.30 < r < .63$, and the PCL-R Antisocial factor, $r = .41$, $.22 < r < .57$.

PPI Criterion reference

Some previous researchers have used a mean or median total PPI score as a cut-off for categorizing individuals as psychopathic versus non-psychopathic (Vaughn, Howard and Delisi 2008). We introduce the mean score of the healthy participants as a criterion reference score for psychopathy. The criterion reference score allows researches to compare individual test scores with a standard criterion reference (to generate an individual z-score indicating an individual's position with respect to the group) instead of a random study sample.

Constructing the PPI-SF

Scalability We first conducted the dimensionality analysis on one random half of the data and replicated the results on the second random half to evaluate the influence of chance capitalization. Both analyses yielded the same results; therefore only the results obtained in the first data set are discussed. Dimensionality analysis on the 30 items for Machiavellian Egocentricity suggested a medium 16-item scale ($H = .44$). Dimensionality analysis for Social Potency yielded an acceptable, weak 17-item scale ($H = .35$). For Coldheartedness we found a medium 9-item solution ($H = .42$). Dimensionality analysis for Carefree Nonplanfulness yielded

a medium 13-item scale ($H = .44$). For Fearlessness a medium 12-item scale ($H = .43$) was accepted. For Blame Externalization a strong 10-item scale ($H = .50$) was accepted. For Impulsive Nonconformity we found a weak 10-item scale ($H = .33$). For Stress Immunity a medium 7-item scale ($H = .49$) was accepted. Contrary to the results of Benning et al. (2003), dimensionality analysis of the PPI two-factor structure did not yield useful results[2]. Hence, our stable scaling results were the point of departure for further analyses pursued here.

Item Location Items exhibiting a mean score of approximately 2.5 reflect an average of the psychopathic trait referred to in the item text. For Machiavellian Egocentricity, an average item was "I'm good at flattering important people when it's useful to do so." (Item #66). An average item for Social Potency was "If I really wanted to, I could convince most people of just about anything." (Item #49). An average Coldheartedness item was "Seeing a poor or homeless person walking the streets at night would really break my heart." (Item #69). An average Carefree Nonplanfulness item was "I usually strive to be the best at whatever I do." (Item #24). An average item for Fearlessness was "When my life becomes boring I like to take some chances to make things interesting." (Item #45). An average item for Blame Externalization was "I've been the victim of a lot of bad luck in my life." (Item #44). Within the Impulsive Nonconformity scale, an average item was "I would enjoy hitch-hiking my way across the United States with no prearranged plans." (Item #28). An average item for Stress Immunity was "I am easily 'rattled' at critical moments." (Item #63).

---

[2] Results can be obtained on request from the first author.

Item Discrimination For Impulsive Nonconformity, item scalability was weak (H_j in the range .3-.4); for Machiavellian Egocentricity, Social Potency, and Stress Immunity, weak to medium (H_j: .3-.5); for Fearlessness, weak to strong (H_j: .3-.56); and for Coldheartedness, Carefree Nonplanfulness, and Blame Externalization, medium to strong (H_j: .4-.58).

Convergent Validity PPI-SF total score correlated significant with PCL-R total score, $r=.42$ [.24<r<.58], the PCL-R Lifestyle factor, $r=.48$ [.30<r<.63], and the PCL-R Antisocial factor, $r=.40$ [.21<r<.56].

PPI-SF Criterion Reference Table 2 shows criterion reference scores for the PPI-SF ($n=60$). We used the mean score of the healthy comparison participants as a criterion reference score for psychopathy, making it possible to produce an individual z-score indicating the individual's position with respect to the group.

Discussion

The aim of this first study was threefold: first, to examine the reliability and validity of the PPI; second, to develop criterion reference scores; and third, to design a novel PPI Short Form (PPI-SF) using more sophisticated psychometric methods than were used to develop the PPI or PPI-R. Results were consistent with previous research on the original PPI (Ross et al. 2009) in showing satisfactory reliability and broadly satisfactory validity for the PPI. The results revealed a significant difference between forensic and healthy participants, with higher total score for the former group. Unexpectedly, however, we did not find a significant difference between the forensic and healthy controls for PPI Machiavellian Egocentricity, Social Potency, Carefree Nonplanfulness, or Coldheartedness. Originally, the PPI was developed for measuring psychopathic personality characteristics within healthy control populations and does not explicitly assess criminal behavior. Given this purpose, one might expect healthy controls to score high on certain subscales, especially those measuring somewhat adaptive functioning (e.g., Social Potency). However, PPI scores of healthy controls in the current study did not deviate from those in other healthy control samples (Carlson and Thái 2010; Fecteau et al. 2008). Most important, PPI total scores did not show significant differences between the healthy controls from the current study and the comparison healthy control samples used in other studies (Carlson and Thái 2010; Fecteau et al. 2008). We therefore recommend interpreting individual subscale scores only in light of the PPI total score and the criterion reference score.

Moreover, the current study replicated earlier findings (Poythress et al. 1998) concerning a moderate to high correlation between PPI and PCL-R total scores. Further, PPI data were used to construct a less time-consuming screening instrument (PPI-SF) and corresponding criterion reference scores. MSA results for item scalability, item popularity, and item discrimination showed a differentiation in scales with weak (Impulsive Nonconformity and Social Potency) and strong scalability (Coldheartedness, Blame Externalization, and Carefree Nonplanfulness). In addition to the PPI results, we found significant relationships between PPI-SF and PCL-R total scores and between the Behavioral Lifestyle

Table 2 Group Differences for PPI-SF Total Scale and Subscales Between Forensic and Control Groups; Skewness and Kurtosis, Independent-Samples t-Tests, and Effect Size (d)

	Forensic psychopathic		Forensic Non-psychopathic		Controls		Skewness	Kurtosis	t	d
	M	(SD)	M	(SD)	M	(SD)				
Total score	200.20	(20.56)	171.08	(20.89)	181.30	(24.94)	−0.05	−0.47	8.04***	0.77
Machiavellian Egocentricity	38.13	(8.31)	30.16	(7.81)	33.80	(7.03)	0.00	−0.63	5.09**	0.61
Social Potency	41.87	(10.43)	36.00	(7.79)	39.55	(7.78)	−0.72	0.03	2.39	0.42
Coldheartedness	9.53	(3.90)	6.44	(3.68)	10.90	(5.00)	0.66	0.46	6.62**	0.69
Carefree Nonplanfulness	11.53	(5.29)	10.28	(5.05)	14.00	(4.48)	−0.04	−0.67	3.20*	0.48
Fearlessness	33.67	(4.47)	27.96	(7.90)	30.65	(5.81)	−0.62	−0.17	3.65*	0.52
Blame Externalization	27.93	(4.57)	26.92	(6.37)	16.40	(5.13)	−0.04	−0.74	25.70***	1.37
Impulsive Nonconformity	25.60	(4.88)	21.48	(5.32)	24.10	(5.34)	0.00	−0.26	3.20*	0.48
Stress Immunity	14.73	(3.45)	14.12	(3.93)	14.50	(4.01)	0.16	−.23	0.13	0.10
Deviant Responding	1.60	(.63)	1.68	(0.99)	1.15	(.49)				
Unlikely Virtues	13.60	(2.56)	14.28	(3.12)	13.90	(2.34)				
VRIN	1.60	(4.78)	0.44	(5.25)	2.75	(4.44)				

* $p<.05$, ** $p<.01$, *** $p<.001$

and the Antisocial factors of the PCL-R, which are known for their predictive value for violence, antisocial conduct and criminal recidivism (Leistico et al. 2008; Walters et al. 2008).

This study suffered from three limitations. First, the healthy comparison participants showed a more pronounced tendency to answer in a more socially desirable direction than the forensic participants (see Table 1). We cannot readily explain this result because we had only limited access to background information in the former group; nevertheless, the results may simply reflect higher psychological health among individuals in this group. Second, PCL-R scores were available only for a subset of the full sample. Because the PCL-R still is the most commonly used measure to assess psychopathy, the PPI-SF validation study (Study 2) included a PCL-R assessment for all participants as a standard procedure. Third, during the data collection for the current study, Lilienfeld and Widows (2005) published a shortened (154 item) revised version of the PPI, the PPI-Revised (PPI-R). The PPI-R excluded items that were problematic psychometrically as well as items that were difficult to read or culturally specific. In contrast, in the present study, a sophisticated psychometric analysis was used to exclude items from the PPI-SF, resulting in the shortest questionnaire justifiable. Given the need for a less time consuming screening device, the PPI-SF may be the best alternative. Moreover, as of this writing the original PPI continues to be used extensively in research and clinical settings. Therefore, we based the analytic approach on the original PPI. Clearly, in further research the current approach should be extended to the PPI-R as well.

Study 2

The second study had two major goals. First, we examined the reliability and the validity of the newly developed PPI-SF. Second, we developed criterion reference scores for this PPI-SF.

Method

Participants

PPI-SF and PCL-R data were collected from 60 male participants (15 forensic psychopathic, 25 forensic non-psychopathic, and 20 healthy controls). The Ethical Committee of Maastricht University and the Ministry of Justice approved the research protocol. A psychologist recruited forensic participants at Penitentiary Institution 'De Geerhorst' Sittard. Healthy comparison participants were recruited at Maastricht University. All participants cooperated voluntarily

and received written and oral instructions. Forensic participants were informed that participation was not related to treatment or criteria for parole. All participants were free to withdraw from the study at any time. Participants were in majority Caucasian (87 % of forensic psychopathic, 92 % of forensic non-psychopathic, and 100 % of healthy controls) and ranged in age from 19 to 59 years (M=32.4, SD=10.3). Forensic participants had a mean age of 32.4 years (SD=10.3), and healthy controls had a mean age of 24.7 years, SD=3.4, $t(53)$=4.3, p=.00, d=1.2[3]. Twenty-five percent of the forensic participants were convicted for (attempted) manslaughter or murder, 30 % for bodily harm, 23 % for property crime with violence, 13 % for property crime, 8 % for sexual offences, and 2 % were convicted for other crimes. PCL-R data were collected from all participants. Supplementary institutional files and collateral information were available for the forensic participants.

Measurement Instruments

Psychopathic Personality Inventory-Short Form ([PPI-SF] Lilienfeld and Andrews 1996) The 100-item PPI-SF from Study 1 (see above) was used.

Psychopathy Checklist-Revised ([PCL-R] Hare 1991, 2003) The PCL-R (see Study 1) was used. All PCL-R interviews were videotaped and scored by two supervised, trained masters' students. PCL-R assessments in a subsample (n=15) of the current forensic sample were double-scored with a consensus rating arrived at by discussion. Inter-rater reliability, expressed by the intraclass-correlation coefficient (ICC), equaled .98 for the total PCL-R score (Landis & Koch, 1977). ICC for the PCL-R factors ranged from .90 (Interpersonal factor) to .99 (Antisocial factor).

Reactive-Proactive Aggression Questionnaire ([RPQ] Raine et al. 2006) The RPQ consists of 23 items rated on a 3-point scale (0=*never* and 2=*often*). The RPQ contains two subscales: a Proactive subscale containing 12 items assessing predatory aggression and a Reactive subscale consisting of 11 items assessing impulsive aggression. Earlier studies showed good reliability, good convergent and discriminant validity for RPQ scores with coefficient alpha for the Reactive subscale ranging from .81 to .86, and for the Proactive subscale scores ranging from .84 to .86 (Raine et al. 2006).

The Aggression Questionnaire ([AQ] Buss and Perry 1992) The AQ is a widely used self-report measure containing

---

[3] Age-corrected results were similar to the reported results (with one exception: no significant group difference was found in the age-corrected PPI-SF analysis in Carefree Nonplanfullness). Both results can be obtained on request from the first author.

29 items that are answered on a 5-point scale (0=*definitely disagree* and 4=*definitely agree*). The AQ has a four-factor structure comprising physical aggression, verbal aggression, hostility, and anger. The test-retest reliability of the AQ ranges between .72 and .80 across various intervals (Hornsveld, Muris, Kraaimaat, & Meesters, 2009). The AQ also displays strong convergent validity, as demonstrated by high correlations with other measures of aggression (Buss and Perry 1992).

The Behavioral Inhibition System and Behavioral Activation System scales ([BIS-BAS] Carver and White 1994) The BIS-BAS is a self-report measure consisting of 20 items, rated on a 4-point scale (1=*strongly agree* and 4=*strongly disagree*). The BIS-BAS contains four subscales: the BIS scale, assessing sensitivity to punishment cues; the BAS-Reward Responsiveness scale, assessing sensitivity to rewards and reward cues; the BAS-Fun Seeking scale, assessing motivation to achieve goals; and the BAS-Drive scale, assessing the willingness to rewarding stimuli without reflection. The BIS-BAS scales show moderate to good internal consistency reliability, with alphas ranging from .59 to .79 (Fayers and Machin 2007).

Statistical Analysis

The distribution of the total PPI-SF scores was approximately normal (Table 2). Internal consistency of the PPI-SF was estimated by means of coefficient alpha. One-way multivariate analysis of variance (MANOVA) was used to evaluate group differences among forensic psychopathic, forensic non-psychopathic, and healthy participants (based on PCL-R assessment with a cut-off score of 26[4]). Convergent validity was investigated using correlations between PPI-SF and aggression measures, as various studies demonstrated increased levels of aggression and insensitivity to punishment cues in psychopathy (Uzieblo et al. 2007), as well as with the BIS scale of the BIS-BAS, which measures sensitivity to punishment cues. The validity of the PPI-SF was further investigated using logistic regression of the target variable PCL-R cut-off score[4] on the PPI-SF total score.

Results

Reliability of the PPI-SF

In the total sample, coefficient alpha for the complete PPI-SF equalled .86. For the PPI-SF subscales, alphas ranged

---

[4] The PCL-R cut-off score of 26 was used as a dichotomous target variable categorizing individuals as psychopathic (PCL-R total score≥26) versus non-psychopathic (PCL-R total score<26). Recently, a lower cut-off score of 26 was suggested for psychopathy in Europe (Cooke 1995; Grann et al. 1998).

from .64 (Stress Immunity) to .87 (Blame Externalization). Alpha for the complete PPI-SF was .85 for forensic participants and .89 for the controls. For forensic participants, PPI-SF subscales exhibited alphas ranging from .57 (Stress Immunity) to .82 (Machiavellian Egocentricity and Social Potency). For controls, alphas ranged from .65 (Carefree Nonplanfulness) to .85 (Blame Externalization).

Validity of the PPI-SF

Group Differences The results showed predicted and significant group differences[3] with respect to PPI-SF total scores on the one hand, and the PPI-SF-subscales of Machiavellian Egocentricity, Coldheartedness, Carefree Nonplanfulness, Fearlessness, Blame Externalization, and Impulsive Nonconformity, on the other (Table 2). No group differences were found for the PPI-SF-subscales of Social Potency and Stress Immunity.

Bonferroni-corrected post-hoc tests[3] revealed a statistically significant group difference for the PPI-SF-total score, for the forensic psychopathic versus forensic non-psychopathic participants, $t(56)=4.01$, $p<.01$, $d=.1.07$, and for forensic psychopathic versus healthy participants, $t(56)=2.49$, $p<.05$, $d=.67$. Moreover, forensic psychopathic participants scored significantly higher than forensic non-psychopathic participants on Machiavellian Egocentricity, $t(56)=3.18$, $p<.01$, $d=.85$, and Fearlessness, $t(56)=2.68$, $p<.05$, $d=.72$. Both forensic groups scored significantly higher than healthy participants on Blame Externalization (forensic psychopathic participants, $t(56)=6.06$, $p<.001$, $d=1.62$, and forensic non-psychopathic participants, $t(56)=6.30$, $p<.001$, $d=1.68$). Healthy controls did not differ significantly from psychopathic participants on Carefree Nonplanfulness and Coldheartedness.

Convergent Validity PPI-SF total score correlated significantly with the PCL-R total score, $r=.39$, $.15<r<.59$[3] and all PCL-R factors (Factor 1, $r=.32$ [$.07<r<.53$], Factor 2, $r=.40$ [$.16<r<.59$], Factor 3, $r=.32$ [$.07<r<.53$], and Factor 4, $r=.30$ [$.05<r<.51$]).

Correlations between the PPI-SF total score and the AQ were significant for the AQ total score ($r=.47$ [$.25<r<.65$]) and three AQ subscale scores: Physical Aggression ($r=.53$ [$.32<r<.69$]); Verbal Aggression ($r=.50$ [$.28<r<.67$]); and Anger ($r=.40$ [$.16<r<.59$]), but for Hostility the correlation was not significant.

Correlations between the total PPI-SF score and the BAS subscales of the BIS-BAS were not significant. Correlations between the PPI-SF total score and the RPQ yielded significant results for the RPQ total scale, $r=.63$ [$.45<r<.76$], the Proactive aggression scale, $r=.61$ [$.42<r<.75$], and the Reactive aggression scale, $r=.53$ [$.32<r<.69$].

PPI-SF and the BIS scale of the BIS-BAS correlated significantly, $r=-.39$ [$-.59<r<-.15$][3].

Results of the logistic regression of PCL-R cut-off score[4] on the PPI-SF total score showed that the PPI-SF displays criterion-related validity for predicting psychopathic personality traits, $B=.08$, $SE(B)=.03$, $\beta=8.83$, $p<.01$, Nagelkerke $R^2=.45$.

Discussion

The aim of the second study was to investigate the reliability and the validity of the PPI-SF and develop criterion reference scores. Results were broadly consistent with previous research on the PPI (Ross et al. 2009) and with those of the first study, revealing adequate reliability and validity for the PPI-SF. First, analyses examining group differences replicated Study 1 by showing a significantly higher PPI-SF total score for the forensic psychopathic participants compared with the other groups. Moreover, forensic participants obtained significantly higher scores on the PPI-SF subscales of Machiavellian Egocentricity, Fearlessness and Blame Externalization, but unexpectedly, healthy participants scored significantly higher than forensic non-psychopathic participants on Coldheartedness and Carefree Nonplanfulness. Item-loading difficulties (cross loadings) with these content scales and possible elimination of the Coldheartedness scale have been discussed before (Neumann et al. 2008). We advise interpreting individual subscale scores only in light of the PPI total and PPI criterion reference score results. Second, results replicated earlier findings (Poythress et al. 1998) and the results of Study 1, showing a significant correlation between PPI-SF and PCL-R scores. Moreover, the PPI-SF was associated with total PCL-R scores. Furthermore, consistent with earlier findings (Hare 2003; Raine et al. 2006), we found a strong relation between psychopathy and aggression and between psychopathy and self-reported insensitivity to cues of punishment (Uzieblo et al. 2007). However, we did not find a significant relation with hostility, which contrasts with earlier research (e.g., Newman et al. 2005).

The reasons for the unanticipated group differences between forensic non-psychopathic participants and other participants on Coldheartedness and Carefree Nonplanfulness are unclear. These findings could raise questions regarding the validity of these scales; alternatively, they might be related to individual prospects for release and inmates' personal situation, as all forensic individuals were held in custody in a Penitentiary Institution in anticipation of a possible conviction. Further research is needed to replicate these unanticipated findings in other samples.

As noted earlier, a 56-item short form of the PPI (PPI-SF) has been constructed. (Lilienfeld and Hess 2001). Nevertheless, formal criterion reference scores for this measure have not been developed. Moreover, the original short form of the

PPI was developed by selecting items with high factor loadings, which may result in narrower scales than the more methodologically sophisticated approach used here (Smith et al. 2000). Therefore, we emphasize the importance of a well-validated PPI-SF as presented in the current article.

Summary and Concluding Discussion

We found that the PPI is a valid instrument useful in forensic and non-forensic settings for measuring psychopathic personality characteristics with respect to criterion reference scores. For practical settings, in which the PPI is time consuming, this study presents a valid screening device with corresponding criterion reference scores. The PPI-SF may be helpful in forensic and non-forensic populations for detecting psychopathic personality traits. The PPI-SF may also be of interest for risk assessment, as the PPI-SF showed a significant relationship with the PCL-R and aggression. A general screening for psychopathy may be important in many forensic settings given the well-documented implications of this construct for criminal recidivism (Hare 2003).

References

Benning, S. D., Patrick, C. J., Hicks, B. M., Blonigen, D. M., & Krueger, R. F. (2003). Factor structure of the psychopathic personality inventory: validity and implications for clinical assessment. *Psychological Assessment, 15,* 340–350.

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology, 63,* 452–459.

Carlson, S. R., & Thái, S. (2010). ERPs on a continuous performance task and self-reported psychopathic traits: P3 and CNV augmentation are associated with fearless dominance. *Biological Psychology, 85,* 318–330.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *Journal of Personality and Social Psychology, 67,* 319–333.

Cooke, D. J. (1995). Psychopathic disturbance in the Scottish prison population: cross-cultural generalizability of the hare psychopathy checklist. *Psychology, Crime and Law, 2,* 101–118.

Emons, W. H. M., Sijtsma, K., & Pedersen, S. S. (2012). Dimensionality of the Hospital Anxiety and Depression Scale (HADS) in cardiac patients: comparison of mokken scale analysis and factor analysis. *Assessment, 19,* 337–353.

Fayers, P., & Machin, D. (2007). *Quality of life: the assessment, analysis and interpretation of patient-reported outcomes* (2nd ed.). West-Sussex: Wiley Franken, I. H., & Muris, P. (2006). BIS/BAS personality characteristics and college students' substance use. *Personality and Individual Differences, 40,* 1497–1503.

Fecteau, S., Pascual-Leone, A., & Théoret, H. (2008). Psychopathy and the mirror neuron system preliminary findings from a non psychiatric sample. *Psychiatry Research, 160*, 137–144.

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.

Hare, R. D. (1991). *Manual of the Psychopathic Checklist-Revised (PCL-R)*. North Tonawanda: Multi-Health Systems.

Hare, R. D. (2003). *Manual for the hare psychopathy checklist-revises* (2nd ed.). Toronto: Multi-Health Systems.

Hildebrand, M. (2008). Psychopathy checklist-revised en psychopathy checklist: youth version. In T. Giesbrecht, C. de Ruiter, & M. Jelicic (Eds.), *Forensisch psychodiagnostisch gereedschap*. Amsterdam: Harcourt.

Hornsveld, R. H. J., Muris, P., Kraaimaat, F. W., & Meesters, C. (2009). The Aggression Questionnaire in Dutch violent forensic psychiatric patients and secondary vocational students. *Assessment*, 16, 181-192.

Landis, J. R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.

Leistico, A.-M. R., Salekin, R. T., DeCoster, J., & Rogers, R. (2008). A large-scale meta-analysis relating the hare measures of psychopathy to antisocial conduct. *Law and Human Behavior, 32*, 28–45.

Lilienfeld, S. O., & Andrews, B. P. (1996). Development and preliminary validation of a self-report measure of psychopathic personality traits in noncriminal populations. *Journal of Personality Assessment, 66*, 488–524.

Lilienfeld, S. O., & Hess, T. H. (2001). Psychopathic personality traits and somatization: Sex differences and the mediating role of negative emotionality. *Journal of Psychopathology and Behavioral Assessment, 23*, 11–24.

Lilienfeld, S. O., & Widows, M. R. (2005). *The psychopathic personality inventory-revised*. Lutz, Florida: Psychological Assessment Resources.

Mokken, R. J. (1971). *A Theory and Procedure of Scale Analysis*. Berlin: De Gruyter.

Molenaar, I. W., & Sijtsma, K. (2000). *MSP5 for windows. Version 5.0*. Groningen: Iec ProGAMMA.

Morrissey, C., Hogue, T., Mooney, P., Lindsay, W. R., Steptoe, L., Taylor, J., & Johnston, S. (2005). Applicability, reliability, and validity of the psychopathy checklist – revised in offenders with intellectual disabilities: some initial findings. *International Journal of Forensic Mental Health, 4*, 207–220.

Neumann, C. S., Malterer, M. B., & Newman, J. P. (2008). Factor structure of the psychopathic personality inventory (PPI): findings from a large incarcerated sample. *Psychological Assessment, 20*, 169–174.

Newman, J. P., MacCoon, D. G., Vaughn, L. J., & Sadeh, N. (2005). Validating a distinction between primary and secondary psychopathy with measures of Gray's BIS and BAS constructs. *Journal of Abnormal Psychology, 114*, 319–323.

Poythress, N. G., Edens, J. F., & Lilienfeld, S. O. (1998). Criterion-related validity of the psychopathic personality inventory in a prison sample. *Psychological Assessment, 10*, 426–430.

Raine, A., Dodge, K., Loever, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., & Liu, J. (2006). The reactive-proactive agression questionaire: differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior, 32*, 159–171.

Ross, S. R., Benning, S. D., Patrick, C. J., Thompson, A., & Thurston, A. (2009). Factors of the psychopathic personality inventory. Criterion-related validity and relationship to the factors of the psychopathic personality inventory. *Assessment, 16*, 71–87.

Sijtsma, K., & Molenaar, I. W. (2002). *Introduction to nonparametric item response theory*. Thousand Oaks: Sage.

Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment, 12*, 102–111.

Uzieblo, K., Verschuere, B., & Crombez, G. (2007). The psychopathic personality inventory: construct validation of the two-factor structure. *Personality and Individual Differences, 43*, 657–667.

Van Ginkel, J. R., & Van der Ark, L. A. (2005). SPSS syntax for missing value imputation in test and questionnaire data. *Applied Psychological Measurement, 29*, 152–153.

Vaughn, M. G., Howard, M. O., & Delisi, M. (2008). Psychopathic personality traits and delinquent careers: an empirical examination. *International Journal of Law and Psychiatry, 31*, 407–416.

Vaughn, M. G., Newhill, C. E., DeLisi, M., Beaver, K. M., & Howard, M. O. (2008). An investigation of psychopathic features among delinquent girls: violence, theft, and drug abuse. *Youth Violence and Juvenile Justice, 6*, 240–255.

Verschuere, B., Crombez, G., Koster, E. H. W., & Clercq, D. (2007). Antisociality, underarousal, and the validity of the concealed information polygraph test. *Biological Psychology, 74*, 309–318.

Walters, G. D., Knight, R. A., Grann, M., & Dahle, K. P. (2008). Incremental validity of the psychopathy checklist facet scores: predicting release outcome in six samples. *Journal of Abnormal Psychology, 117*, 396–405.