Violence Risk Assessment and Facet 4 of the Psychopathy Checklist: Predicting Institutional and Community Aggression in Two Forensic Samples

Glenn D. Walters[1] and Kirk Heilbrun[2]

Abstract

The Psychopathy Checklist and Psychopathy Checklist–Revised (PCL/PCL-R) were used to predict institutional aggression and community violence in two groups of forensic patients. Results showed that Facet 4 (Antisocial) of the PCL/PCL-R or one of its parcels consistently achieved incremental validity relative to the first three facets, whereas the first three facets failed to achieve incremental validity relative to Facet 4. One of the two Facet 4 parcels, Parcel G (General Acting Out), was the only PCL-R measure to consistently achieve success in classifying individual cases using the receiver operating characteristic approach. These findings suggest that Facet 4 and its parcels may play a role in violence risk assessment, although the generalizability of these findings to various forensic settings and contexts requires further study.

Keywords

Psychopathy Checklist, Facet 4, Parcel G, forensic patients, violence risk assessment

Violence risk in forensic patients has been assessed with a variety of different procedures, including the Psychopathy Checklist (PCL; Hare, 1980) and Psychopathy Checklist–Revised (PCL-R; Hare, 2003). Hare (1998) notes that the PCL/PCL-R was not designed for risk assessment even though it is capable of predicting violent recidivism at a level commensurate with actual risk assessment procedures. Studies conducted on institutional (Buffington-Vollum, Edens, Johnson, & Johnson, 2002; Kroner & Mills, 2001) and community (Glover, Nicholson, Hemmati, Bernfeld, & Quinsey, 2002; Grann, Långström, Tengström, & Kullgren, 1999) aggression indicate that the PCL-R may be predictive of both. However, the relationship between PCL-measured psychopathy and subsequent community violence is more robust than the relationship between psychopathy and institutional misconduct (Edens, Petrila, & Buffington-Vollum, 2001). In fact, the aspect of institutional misconduct emphasized in many of the studies on the topic bears little resemblance to the physical violence that is of primary concern to clinicians and policy makers. In the Buffington-Vollum et al. (2002) study, for instance, the PCL-R predicted verbal institutional aggression but not physical institutional aggression, and in the Kroner and Mills (2001) study, the PCL-R correlated only weakly with "major misconduct," a category that included rioting, drug/alcohol abuse, and refusing a direct order. Questions have also been raised about the

PCL-R's ability to predict serious assaultive behavior during confinement (Edens, Buffington-Vollum, Keilen, Roskamp, & Anthony, 2005).

It is becoming increasingly apparent that not all aspects of the PCL-R are equally effective in predicting violence. In a meta-analysis comparing the predictive efficacy of the two PCL-R factor scores, Walters (2003) ascertained that Factor 2 (Chronically Unstable Antisocial Lifestyle) was a significantly better predictor of institutional misconduct and recidivism than Factor 1 (Selfish, Callous, and Remorseless Use of Others). More recently, Walters, Knight, Grann, and Dahle (2008) discovered that one of the two facets that comprises Factor 2 of the PCL-R, Facet 4 (Antisocial), predicted violent and general recidivism above and beyond the first three facets, whereas the first three facets generally failed to predict recidivism above and beyond Facet 4. The Walters et al. (2008) study also revealed that of the two Facet 4 parcel scores, Parcel H (Criminality) was a stronger predictor of general recidivism than Parcel G (General Acting

[1]Federal Correctional Institution Schuylkill, Minersville, PA, USA
[2]Drexel University, Philadelphia, PA, USA

Corresponding Author:
Glenn D. Walters, Psychology Services, FCI-Schuylkill,
PO Box 700, Minersville, PA 17954-0700, USA
Email: gwalters@bop.gov

Out), whereas the parcels were equally effective in predicting aggressive recidivism. Further research is required to determine whether Parcels G and H are equally effective in predicting institutional aggression and community violence and whether dynamic items on the PCL-R (e.g., poor behavioral controls) are capable of predicting these outcomes.

Analyzing two popular risk assessment procedures, Mills, Kroner, and Hemmati (2007) discerned that several of the individual items on the Historical-Clinical-Risk–20 (HCR-20; Webster, Eaves, Douglas, & Wintrup, 1995) and Violence Risk Appraisal Guide (VRAG; Webster, Harris, Rice, Cormier, & Quinsey, 1994) failed to discriminate between violent recidivists and nonrecidivists. From their findings, Mills et al. (2007) concluded that retaining non-differentiating items can compromise the validity and utility of a risk assessment procedure such as the HCR-20 or VRAG. They consequently recommend deleting nondiscriminating items and recalibrating others as part of the cross-validation process. The PCL-R total score is included as an item on both the HCR-20 and VRAG; in the Mills et al. (2007) study, the VRAG PCL-R item achieved significance whereas the HCR-20 PCL-R item did not. One explanation for this discrepancy in overall outcome is that by using different cutting scores and dividing the PCL-R total score into different numbers of categories, three in the case of the HCR-20 and six in the case of the VRAG, the two procedures produced divergent results. Part of the problem, however, may be with the PCL-R total score itself. If only one of the four components of the PCL-R is useful in predicting violence, then the other three components may actually impair the predictive efficacy of the PCL-R total score. Under such circumstances, it may be better to replace or supplement the total PCL-R score with Facet 4 or one of its parcels.

For a scale to serve as an effective predictor of violence risk it must contain items that are capable of maximally discriminating between those who will and those who will not engage in future violence (Mills et al., 2007). In addition to items that are maximally discriminating, an effective risk assessment measure should include items that possess incremental validity relative to other actual and potential predictors of violence (Conroy & Murrie, 2007). Results from a recent study by Edens, Skeem, and Douglas (2006) showed that the screening version of the Psychopathy Checklist (PCL:SV; Hart, Cox, & Hare, 1995) possessed incremental validity relative to the remaining portion of the VRAG but that the remaining portion of the VRAG did not possess incremental validity relative to the PCL:SV. Third, a violence risk measure is most flexible when composed of both static (historical) and dynamic (changeable) items. In this way, the measure provides a balance between the interdependent goals of prediction and risk management (Heilbrun, 1997). Although the PCL-R is generally considered a stable personality measure, certain items may be more dynamic than other items. In the current investigation, the incremental validity and classification accuracy of Facet 4 and its two parcel scores will be evaluated against the first three facets. It is hypothesized that (a) the PCL-R Facet 4 score and its two parcels (G and H) will achieve incremental validity relative to the first three PCL-R facets, but the first three facets will not achieve incremental validity relative to Facet 4 or its two parcels; and (b) the PCL-R Facet 4 score and its two parcels (G and H) will accurately classify a significant portion of individual cases as high or low risk using the receiver operating characteristic (ROC) approach.

Method

Participants

Sample 1 for this study was composed of 216 male forensic patients admitted to the Florida State Hospital between 1984 and 1990 after being judged incompetent to stand trial, not guilty by reason of insanity, or dangerous to self or others as part of a civil commitment proceeding (Heilbrun et al., 1998). Participants in Sample 1 ranged in age from 17 to 86 years ($M = 31.72$, $SD = 10.01$), with 119 (55.1%) of the participants being White and 97 (44.9%) of the participants being Black. Sample 2 comprised 230 male inmates housed at the United States Penitentiary in Atlanta, Georgia, between 1991 and 2000. All participants in the second sample had been referred by the federal courts for a pretrial psychiatric evaluation sometime between 1991 and 2000 (Walters & Duncan, 2005; Walters, Geyer, & Duncan, 2003). This second sample ranged in age from 19 to 64 years ($M = 37.00$, $SD = 10.02$) and had between 1 and 20 years of education ($M = 10.20$, $SD = 2.78$). Ethnically, the sample was composed of 115 (49.6%) White inmates, 103 (45.4%) Black inmates, 19 (3.9%) Hispanic inmates, and 3 (1.3%) Asian/Native American inmates. None of the participants in either Sample 1 or Sample 2 were included in the previous Walters et al. (2008) incremental validity analyses.

Measure

The PCL (Hare, 1980) is a 22-item rating scale designed to assess psychopathy. Two items (previous diagnosis as a psychopath and antisocial behavior not due to alcohol intoxication) were subsequently dropped from the scale, one item ("irresponsibility as a parent") was made more general ("irresponsibility"), and the titles of 10 items were changed slightly without altering the underlying trait or behavior to create the 20-item PCL-R (Hare, 2003). Each item on the PCL/PCL-R is rated on a 3-point scale: 0 = *does not apply*, 1 = *may apply or in some respects applies*, 2 = *does apply*. Accordingly, the PCL generates a maximum total score of 44, and the PCL-R generates a maximum

Table 1. Descriptive Statistics and Correlations for PCL Facet Scores and Parcels G and H in Sample 1

					Correlations				
PCL Component	Range	M	SD	α	Facet 2	Facet 3	Facet 4	Parcel G	Parcel H
Facet 1 (Interpersonal)	0-8	3.49	1.94	.71	.46**	.28**	.29**	.18*	.31**
Facet 2 (Affective)	0-8	5.03	1.86	.74		.34**	.30**	.25**	.29**
Facet 3 (Lifestyle)	0-10	6.09	1.97	.63			.45**	.36**	.43**
Facet 4 (Antisocial)	0-10	5.46	2.26	.68				.85**	.94**
Parcel G (General Acting Out)	0-4	2.32	1.01	.18					.61**
Parcel H (Criminality)	0-6	3.14	1.51	.55					

Note: Range = low to high score; M = mean; SD = standard deviation; α = Cronbach alpha coefficient. N = 216.
*p < .01. **p < .001.

Table 2. Descriptive Statistics and Correlations for PCL-R Facet Scores and Parcels G and H in Sample 2

					Correlations				
PCL-R Component	Range	M	SD	α	Facet 2	Facet 3	Facet 4	Parcel G	Parcel H
Facet 1 (Interpersonal)	0-8	3.93	2.57	.86	.80*	.60*	.51*	.52*	.43*
Facet 2 (Affective)	0-8	5.44	2.48	.90		.70*	.52*	.54*	.42*
Facet 3 (Lifestyle)	0-10	6.80	2.28	.79			.64*	.66*	.53*
Facet 4 (Antisocial)	0-10	4.90	2.70	.69				.83*	.94*
Parcel G (General Acting Out)	0-4	2.43	1.13	.49					.60*
Parcel H (Criminality)	0-6	2.46	1.86	.50					

Note: Range = low to high score; M = mean; SD = standard deviation; α = Cronbach alpha coefficient. N = 230.
*p < .001.

total score of 40. Both measures can be organized into two factor scores and four facet scores. There are no differences between the PCL and PCL-R with respect to the items included in the facet scores for each procedure.

Factor 1 (Selfish, Callous, and Remorseless Use of Others) of the PCL-R can be subdivided into two 4-item facets: Facets 1 (Interpersonal: glib/superficial, grandiose self-worth, pathological lying, conning/manipulative) and 2 (Affective: lack of remorse or guilt, shallow affect, callous/lacks empathy, failure to accept responsibility). Factor 2 (Chronically Unstable and Antisocial Lifestyle) can be subdivided into two 5-item facets: Facets 3 (Lifestyle: stimulation seeking/boredom proneness, impulsivity, irresponsibility, parasitic orientation, lack of realistic goals) and 4 (Antisocial: poor behavioral controls, early behavior problems, juvenile delinquency, revocation of conditional release, criminal versatility). Facet 4 is divided further into parcels: Parcel G (General Acting Out: poor behavioral controls, early behavior problems) and Parcel H (Criminality: juvenile delinquency, revocation of conditional release, criminal versatility).

Reliability and validity data collected on the PCL and PCL-R confirm the stability and predictability of the scores obtained with these measures (Hare, 1980, 2003). The inter-rater reliability of the total PCL/PCL-R score was calculated using the intraclass correlation coefficient (ICC). Two trained raters independently scored each participant in Sample 1 from a single interview. The single-measures ICC (absolute agreement) was .88 between raters. One trained rater interviewed and scored each participant in Study 2 with the PCL-R. The interviews were taped, however, and 10 randomly selected interviews were reviewed and scored by three doctoral-level clinical psychology interns trained in the use of the PCL-R. This produced a single measures ICC (absolute agreement) of .81. The ranges, means, standard deviations, alpha coefficients, and intercorrelations of the six PCL/PCL-R measures used in this study are listed in Tables 1 (Sample 1) and 2 (Sample 2).

Procedure

Sample 1 participants were independently rated on the 22-item PCL by two graduate research assistants from a single semistructured interview and a review of records that predated the current hospital admission and the results averaged. These raters were trained to administer and score the PCL by Kirk Heilbrun. PCLs were completed within a month of a patient's admission to the hospital and were administered exclusively for research purposes. As such, informed consent to participate in the research study was obtained from each participant. Item ratings were summed to form a Facet 1 score, Facet 2 score, Facet 3 score, Facet 4

score, Parcel G score, and Parcel H score. An independent group of raters then retrospectively reviewed a participant's hospital chart for evidence of violent incidents (shouting, threats, pushing, or hitting) during the first 3 months and last 3 months of hospitalization. Of the 195 participants with complete hospital data, 75 (38.5%) engaged in one or more episodes of violence during the first 3 or last 3 months of hospitalization. A binomial logistic regression analysis was performed on the dichotomous institutional violence outcome measure (present/absent) using Version 17 of the Statistical Package for the Social Sciences (SPSS, 2008). Follow-up data subsequent to hospital discharge were available for 183 of the participants in Sample 1 and were obtained from records maintained by the Florida Department of Law Enforcement and Federal Bureau of Investigation (FBI). The follow-up period ranged from several months to 6 years ($M = 4.19$ years, $SD = 1.48$ years). Arrests for violent reoffenses (person crimes) were converted to annual rates and the results subjected to negative binomial regression analysis as computed by Version 8.0 of the LIMDEP statistical program (Greene, 2002). Out of 183 released participants, 67 (36.6%) were arrested for a violent offense at least once during the follow-up.

Sample 2 participants were scored on the 20-item PCL-R by a psychologist (Scott Duncan) with extensive training and experience in the use of this procedure. This psychologist also trained and supervised the psychology interns who participated in the interrater reliability analyses previously described. Participants were administered the PCL-R for clinical purposes and so informed consent was not obtained, although an institutional review board did approve the use of these data for research. Item ratings were summed to form a Facet 1 score, Facet 2 score, Facet 3 score, Facet 4 score, Parcel G score, and Parcel H score. The electronic disciplinary records of all participants who remained in federal custody for 2 years after administration of the PCL-R ($N = 185$) were reviewed and aggressive disciplinary infractions (fighting, assault, threatening) recorded. A formal hearing was held and only infractions for which an inmate pled or was found guilty of were included in these analyses. Forty-three of the 185 inmates (23.2%) were found guilty of committing an aggressive infraction (six of which were classified as severe assaults resulting in hospitalization or death) sometime during the 2-year follow-up. The dichotomous outcome (present/absent) was subjected to binomial logistic regression analysis using SPSS 17.0. There were 122 inmates from Sample 2 who had been released from custody during a 1- to 152-month follow-up ($M = 69.32$, $SD = 41.66$). Eleven of these inmates (9.0%) were arrested for one or more violent person crimes (assault, murder, rape, robbery) according to the results of a review of files maintained at the FBI's National Crime Information Center. Because time until first arrest for a violent offense was

available for Sample 2, a stepwise Cox regression analysis was performed on the violent recidivism data with SPSS 17.0. All outcome data in Sample 2 were collected subsequent to administration of the PCL-R.

Results

Incremental Validity Analyses

A binary logistic regression analysis of institutional aggression in Sample 1 where Facets 1, 2, and 3 were entered at Block 1 and either Facet 4 or Parcels G and H were entered at Block 2 was performed. The results revealed a significant Block 1 effect, $\chi^2(3) = 10.61$, $p < .05$, and significant Block 2 effects for Facet 4, $\chi^2(1) = 4.63$, $p < .05$, and Parcels G and H, $\chi^2(2) = 7.76$, $p < .05$. When the order of predictors was reversed, Facets 1, 2, and 3 failed to achieve incremental validity relative to Facet 4, $\chi^2(3) = 5.32$, $p = .15$, or Parcels G and H, $\chi^2(3) = 6.07$, $p = .11$. In the final logistic regression equation for the four-facet model, Facet 4 showed a significant effect, (exp[B] = 1.08 [1.01-1.17],[1] Wald = 4.53, $p < .05$) but Facets 1 (exp[B] = 1.06 [0.97-1.15], Wald = 1.42, $p = .23$), 2 (exp[B] = 1.06 [0.96-1.17], Wald = 1.47, $p = .22$), and 3 (exp[B] = 1.00 [0.91-1.09], Wald = 0.00, $p = .97$) were nonsignificant. In the final equation where Parcels G and H were entered at Block 2, Parcel G (exp[B] = 1.28 [1.05-1.55], Wald = 5.81, $p < .05$) achieved significance, but Parcel H (exp[B] = 0.98 [0.86-1.12], Wald = 0.06, $p = .81$) and Facets 1 (exp[B] = 1.06 [0.97-1.16], Wald = 1.92, $p = .17$), 2 (exp[B] = 1.06 [0.96-1.17], Wald = 1.40, $p = .24$), and 3 (exp[B] = 1.00 [0.92-1.09], Wald = 0.00, $p = .99$) did not.

The violent recidivism data from Sample 1 were subjected to negative binomial regression analysis after it was determined that the violence rate distribution was overdispersed ($t_{obs} = 3.12-3.17$) but not zero inflated (Vuong statistic = 0.81-0.85). A negative binomial regression of the four facet scores revealed a significant effect for Facet 4 ($\beta = .08$ [.03],[2] $t = 2.32$, $p < .05$) but no significant effects for Facets 1 ($\beta = .06$ [.14], $t = 0.43$, $p = .66$), 2 ($\beta = .02$ [.16], $t = 0.12$, $p = .91$), or 3 ($\beta = .05$ [.15], $t = 0.34$, $p = .73$). There were no significant negative binomial regression effects when the first three facets and the two Factor 4 parcels were analyzed ($p > .10$), although the lack of effect for Parcels G and H could have been a consequence of their moderately high correlation ($r = .61$). In fact, when Parcel H was removed from the equation, Parcel G achieved a significant effect ($\beta = .31$ [.14], $t = 2.19$, $p < .05$), and when Parcel G was removed from the equation, Parcel H displayed an effect that approached statistical significance ($\beta = .10$ [.05], $t = 1.93$, $p = .05$). Facet 1, 2, or 3 failed to achieve significance ($p > .10$) in the latter two analyses.

Table 3. Receiver Operating Characteristic (ROC) Results by Sample and Outcome Measure

	Sample 1						Sample 2					
	Institutional Aggression			Violent Recidivism			Institutional Aggression			Violent Recidivism		
Scale	AUC	SE	95% CI	AUC	SE	95% CI	AUC	SE	95% CI	AUC	SE	95% CI
Facet 1	.613**	.042	.530-.695	.561	.043	.476-.646	.532	.049	.436-.627	.499	.062	.377-.621
Facet 2	.589*	.040	.510-.668	.544	.045	.457-.632	.559	049	.464-.654	.569	.079	.415-.723
Facet 3	.570	.042	.489-.652	.556	.044	.471-.642	.569	.050	.470-.667	.596	.076	.448-.744
Facet 4	.626**	.040	.548-.705	.608*	.042	.526-.690	.598	.050	.500-.696	.604	.083	.441-.766
Parcel G	.637**	.040	.558-.716	.597*	.042	.514-.680	.662**	.046	.571-.753	.704*	.078	.551-.858
Parcel H	.597*	.041	.517-.677	.590*	.042	.508-.673	.537	.050	.439-.635	.521	.081	.362-.680

Note: AUC = area under the curve; SE = standard error. 95% CI = asymptotic 95% confidence interval of AUC; Facet 1 = Interpersonal; Facet 2 = Affective; Facet 3 = Lifestyle; Facet 4 = Antisocial; Parcel G = General Acting Out; Parcel H = Criminality. Sample 1 Institutional Violence, $N = 195$; Sample 1 Violent Recidivism, $N = 183$; Sample 2 Institutional Violence, $N = 185$; Sample 2 Violent Recidivism, $N = 122$.
*$p < .05$. **$p < .01$.

A binary logistic regression analysis was performed on aggressive disciplinary reports received by participants in Sample 2. There was no Block 1 effect for the first three facet scores, $\chi^2(3) = 3.11$, $p = .38$, and no Block 2 effect for Facet 4, $\chi^2(1) = 1.74$, $p = .19$. There was, however, a significant Block 2 effect for Parcels G and H, $\chi^2(2) = 12.03$, $p < .01$. In the final logistic regression for the four-facet model, none of the facet scores achieved significance ($p > .10$). In the final logistic regression equation for the three-facet two-parcel model, Parcel G achieved a significant effect (exp[B] = 2.22 [1.39-3.56], Wald = 11.01, $p < .001$), but Facet 1 (exp[B] = 0.89 [0.70-1.13], Wald = 0.91, $p = .34$), Facet 2 (exp[B] = 1.10 [0.83-1.45], Wald = 0.43, $p = .51$), Facet 3 (exp[B] = 0.94 [0.72-1.23], Wald = 0.22, $p = .64$), and Parcel H (exp[B] = 0.86 [0.67-1.09], Wald = 1.58, $p = .21$) failed to achieve significance. When the analyses were restricted to six severely violent infractions (assault leading to hospitalization or death) there were no significant Block 1, $\chi^2(3) = 4.77$, $p = .19$, or Block 2, $\chi^2(1) = 1.87$, $p = .17$, effects in the four-facet model, but there was a significant Block 2 effect, $\chi^2(2) = 6.70$, $p < .05$, for Parcels G and H in the three-facet two-parcel model.

A stepwise Cox regression analysis of violent recidivism in Sample 2 failed to show a significant effect when Facets 1, 2, and 3 were entered at Block 1, $\chi^2(3) = 4.45$, $p = .22$, or when Facet 4 was entered at Block 2, $\chi^2(1) = 0.25$, $p = .62$. A significant Block 2 effect did surface, however, when Parcels G and H were entered at Block 2, $\chi^2(2) = 8.94$, $p < .05$. There was no significant Facet 1, 2, 3, or 4 effect in the final Cox regression equation for the four-facet model ($p > .10$), but there was a significant Parcel G effect (exp[B] = 4.16 [1.50-11.50], Wald = 7.46, $p < .01$) in the final Cox regression equation for the three-facet two-parcel model. In the final equation of the three-facet two-parcel model, Facet 1 (exp[B] = 0.62 [0.39-0.79], Wald = 4.37, $p < .05$) produced a negative predictive effect (i.e., lower rather than higher

scores predicted violent recidivism), whereas Facets 2 (exp[B] = 1.30 [0.82-2.07], Wald = 1.22, $p = .27$) and 3 (exp[B] = 1.11 [0.67-1.82], Wald = 0.16, $p = .69$), and Parcel H (exp[B] = 0.67 [0.43-1.04], Wald = 3.26, $p = .07$) failed to achieve significance.

Although retrospective power analysis is a controversial practice, it is sometimes used to identify analyses that could potentially have produced significant findings had the sample been larger (Thomas, 1997). The largest effect size for each analysis was therefore used to calculate the power of each analysis. In Sample 1, Parcel G achieved .67 power in the analysis of institutional aggression, and Facet 4 achieved .64 power in the analysis of violent recidivism. In Sample 2, Parcel G achieved .91 power in the analysis of institutional aggression and .75 power in the analysis of violent recidivism. Therefore, only one of the four analyses produced a power effect beyond the traditional threshold of .80.

Receiver Operating Characteristic Analyses

ROC analysis denotes the extent to which a measure accurately classifies an outcome into one of two mutually exclusive categories, in this case the presence/absence of institutional aggression and the presence/absence of violent recidivism. The area under the curve (AUC) value delineates the probability that a measure will rank a randomly selected positive outcome (i.e., violence) higher than a randomly selected negative outcome (i.e., nonviolence). The AUC ranges from .500 to 1.00: AUCs between .600 and .699 are typically considered modest, AUCs between .700 and .799 moderate, and AUCs between .800 and 1.000 strong. Table 3 lists the results of the four ROC analyses performed on the two outcome measures in the two samples included in this study. The AUC results for severe institutional assaults leading to hospitalization or death in Sample 2 were slightly

higher than the figures in Table 3; Facet 1 (AUC = .689, *p* = .12), Facet 2 (AUC = .706, *p* = .09), Facet 3 (AUC = .676, *p* = .14), Facet 4 (AUC = .775, *p* < .05), Parcel G (AUC = .833, *p* < .01), and Parcel H (AUC = .683, *p* = .13).

The two items that form Parcel G (poor behavioral controls and early behavior problems) were subjected to ROC analysis to determine whether predictability extends down to the individual items. In Sample 1, poor behavioral controls successfully predicted institutional aggression (AUC = .615, *p* < .01) and approached significance as a predictor of violent recidivism (AUC = .587, *p* = .05). Early behavior problems attained a significant institutional aggression effect (AUC = .588, *p* < .05) but no violent recidivism effect (AUC = .556, *p* = .21). In Sample 2, early behavior problems displayed a significant institutional aggression effect (AUC = .654, *p* < .01) and a significant violent recidivism effect (AUC = .703, *p* < .05). Poor behavioral controls failed to register a significant institutional aggression (AUC = .570, *p* = .16) or violent recidivism (AUC = .586, *p* = .35) effect in Sample 2. Finally, despite the low rate of serious institutional assaults leading to hospitalization or death in Sample 2 (*n* = 6, base rate = 4.9%), early behavior problems successfully predicted this outcome (AUC = .831, *p* < .01), whereas poor behavioral controls did not (AUC = .615, *p* = .34).

Discussion

The current study examined the incremental validity and classification accuracy of the four PCL/PCL-R facet scores and two Facet 4 parcel scores as predictors of institutional aggression and community violence in two different forensic samples. Results for Sample 1 paralleled outcomes obtained by Walters et al. (2008) such that Facet 4 (Antisocial) displayed incremental validity relative to the first three facets, but the first three facets failed to achieve incremental validity relative to Facet 4. Although Sample 2 did not support the incremental validity of any of the four facet scores, one of the two Facet 4 parcels, Parcel G (General Acting Out), attained incremental validity relative to the first three facets and Parcel H (Criminality) in three of the four analyses (institutional aggression in Sample 1, institutional aggression in Sample 2, violent recidivism in Sample 2) and incremental validity relative to the first three facets in the fourth analysis (violent recidivism in Sample 1). In the Walters et al. (2008) investigation, Parcel H outperformed Parcel G in predicting general recidivism, but the two parcels were equally effective in predicting violent recidivism. In the present study, Parcel H performed poorly overall, whereas Parcel G achieved significance in all 10 incremental validity and ROC analyses performed as part of this study. Parcel G, it would seem, may describe a specific domain within psychopathy or a separate construct

altogether that is more strongly related to aggressive outcomes in correctional and forensic contexts than is the broader construct of psychopathy.

Considering this possibility, the current study, unlike the Walters et al. (2008) investigation, examined the classification accuracy of the PCL-R facet scores. It was predicted that Facet 4, Parcel G, and Parcel H but not Facet 1, 2, or 3 would accurately classify individual cases as high or low risk. In support of this hypothesis, Facets 1, 2, and 3 achieved significance in no more than one of four analyses (violent recidivism in Sample 1), whereas Facet 4 achieved significance in two out of four analyses (institutional adjustment and violent recidivism in Sample 1) and Parcel G achieved significance in all four analyses. Despite these modestly encouraging results, there is insufficient evidence at this point in time to conclude that Facet 4 or Parcel G are suitable replacements for the total PCL-R score as items on the HCR-20 or VRAG. The effect was modest overall, and despite being statistically significant in several instances is of unknown clinical utility. There is, nonetheless, preliminary support from both the current study and the previous Walters et al. (2008) investigation to speculate that Facet 4 and Parcel G may serve an important supplemental function in violence risk assessment. Further research is required to determine whether Facet 4 and Parcel G predict institutional aggression and community violence because of their association with psychopathy or for reasons unrelated to their association with psychopathy. Is it that Facet 4 and Parcel G are fundamental aspects of psychopathy with important implications for violence risk assessment or merely a testament to the old adage that the best predictor of future (antisocial) behavior is past (antisocial) behavior.

At least one previous study suggests that the PCL-R is more effective in predicting verbal aggression than it is in predicting physical aggression (Buffington-Vollum et al., 2002). In Sample 1 of the current investigation, there was no way to distinguish between verbal and physical institutional aggression. Consequently, these analyses involved general (verbal + physical) aggression, with verbal aggression being twice as prevalent as physical aggression. The main institutional infraction analyses in Sample 2 also examined general aggression, although a supplemental analysis of serious institutional assaults leading to hospitalization or death was also conducted. In both samples, definitions of recidivism were restricted to person crimes and at least in Sample 2 all individuals who recidivated for verbal aggressive acts (threatening) were removed from the analyses. This reveals the difficulties researchers face when attempting to define aggression and violence (for more information on this issue, see Edens & Douglas, 2006). Aggression was defined in this study as a combination of physical and verbal acts, whereas violence was considered a purely physical act. The significance this distinction has for the

generalizability of the current findings is that we cannot be sure how well some of the results, particularly those pertaining to general institutional aggression, relate to important forensic questions such as a person's level of dangerousness to self or others. In most jurisdictions a person is considered a danger to self or others if he or she poses a threat of physical violence. This definition would seem to require more than simply verbal aggression, thus bringing the generalizability of some of the results obtained in the current study into question.

Although the present study failed to identify a difference in the PCL-R's ability to predict institutional aggression and violent recidivism, prior research notes that the PCL-R may predict community violence better than it predicts institutional aggression (Edens et al., 2001; Kroner & Mills, 2001). One possible explanation for these findings is that Facet 4 items are based more on past community behavior (early behavioral problems, juvenile delinquency, revocation of conditional release, criminal versatility) than on past prison behavior and so correlate better with future community violence than with future prison aggression. Second, aggression in the highly structured prison environment may differ greatly from aggression in the less structured community environment. Third, there may be differences in how aggression is defined or measured in the community versus how it is defined and measured in prison. Kroner and Mills (2001) employed a broad definition of "major misconduct" that featured nonviolent acts such as drug use and refusing an order in their definition of institutional aggression; Buffington-Vollum et al. (2002), on the other hand, discerned that the PCL-R correlated with verbal prison aggression but not physical prison aggression. In the current study, the highest AUC values were obtained when the PCL-R was correlated with severe institutional aggression resulting in a victim's hospitalization or death. Researchers need to consider the severity level of the aggression as well as the environment in which it takes place when determining whether the PCL-R predicts community violence better than institutional aggression or whether this finding is simply an artifact of how prison aggression has been conceptualized, defined, and measured.

Heilbrun (1997) distinguishes between prediction-oriented and management-oriented risk assessment. Whereas the goal of prediction-oriented risk assessment is to accurately predict the probability of a future event's occurrence, management-oriented risk assessment is designed to reduce the odds of a future event's occurrence. Consequently, prediction-oriented risk assessment makes use of static and dynamic risk factors, is insensitive to the changes that occur in risk status, and has minimal implications for intervention. Management-oriented risk assessment, by comparison, relies on dynamic risk factors, is sensitive to risk status change, and has strong implications for intervention. Despite

clear differences between the two methods, Heilbrun (1997) believes that prediction-oriented and management-oriented risk assessment are both equally applicable to a number of legal decision-making contexts. The decision to civilly commit a potentially violent individual with serious mental illness, or an individual completing a sentence following conviction for a sexual offense, is best made using a prediction-oriented approach. However, once the individual has been committed, management-oriented risk assessment is important to guide intervention planning and measure progress. In the current study, Parcel G displayed a consistent ability to predict institutional aggression and community violence. The two PCL-R items that form Parcel G seem to reflect the intercorrelated nature of the prediction-oriented and management-oriented schemes. The static item (early behavioral problems) may be a more reliable predictor of violence (3 out of 4 significant effects) whereas the dynamic item (poor behavioral control) has some predictive utility, but its greater contribution may be in assessing change in a person's risk status.

The findings from this study are consistent with Walters et al.'s (2008) observation that of the four PCL-R facet scores, Facet 4 is the strongest and most incrementally valid predictor of crime-related behavioral outcomes such as institutional aggression and violent recidivism. The robustness of these findings is revealed in the fact that they were obtained in two separate groups of individuals, in different settings, at different times, and in samples not included in the original Walters et al. (2008) six-sample study. In addition, the present study suggests that Parcel G may be the component that is most predictive of violence and the component most useful for violence risk assessment. The predictive power of Parcel G was observed for both institutional aggression and violent recidivism, whether the follow-up period was fixed or variable, and whether the analyses were conducted at the group (logistic regression, negative binomial regression, Cox regression) or individual (ROC) level. The alternation in overall effectiveness for Parcels G and H between the Walters et al. (2008) and current studies may have something to do with the outcome measures used in each study. In the Walters et al. (2008) investigation, Parcel H was a better predictor of general recidivism than Parcel G but the two parcels were equally effective in predicting verbally and physically aggressive recidivism. In the current study, the focus was on aggressive institutional infractions and violent recidivism and as aggression became more exclusively physical Parcel G displayed a growing advantage over Parcel H.

Controversy continues to rage over whether Facet 4 of the PCL-R is a core feature of the psychopathy construct or an entirely different construct altogether. Basing their conclusions on item response theory and other forms of analysis, one group of researchers (Cooke, Michie, &

Skeem, 2007) contend that antisociality (Facet 4) is a consequence rather than a component of psychopathy. Hare and Neumann (2006) counter that Cleckley (1941/1976) included general antisocial behavior in his definition of psychopathy and that Facet 4 is appropriate for inclusion on the PCL-R. Because Facet 4 displays weaker loadings on the psychopathy construct (Cooke et al., 2007) and stronger correlations with violence and aggression (Walters et al., 2008; current study) than the first three facets, it may be that psychopathy is not as strong a predictor of violence and aggression as has traditionally been assumed and that a simple measure of prior criminal or antisocial behavior is a more efficient and practical measure of violence risk than the full PCL-R.

Irrespective of whether Facet 4 is a core element or a consequence of psychopathy, it could clearly be measured better. There is a growing consensus from recent taxometric studies that the latent structure of psychopathy, as measured by the Psychopathy Checklist, is dimensional in nature (Edens, Marcus, Lilienfeld, & Poythress, 2006; Guay, Ruscio, Knight, & Hare, 2007; Walters, Gray, et al., 2007). Indeed, Guay et al. (2007) report that Facet 4 possesses a dimensional latent structure. The similarity between Facet 4 of the PCL-R and measures of antisocial personality disorder are made more salient by studies suggesting that the latent structure of antisocial personality disorder is also dimensional (Marcus, Lilienfeld, Edens, & Poythress, 2006; Walters, Diamond, Magaletta, Geyer, & Duncan, 2007). Because dimensional latent constructs are best assessed by measures that encompass the full range of trait levels (Ruscio, Haslam, & Ruscio, 2006), they require a larger number and greater variety of items than the 11-point Facet 4 scale, 5-point Parcel G scale, or 7-point Parcel H scale. Development and expansion of the Facet 4 items for use in risk assessment may therefore be an important next step in research on the PCL-R and violence risk assessment.

Lynam, Hoyle, and Newman (2006) have discussed the perils of partialling variance in multivariate regression analyses such as the logistic regression, negative binomial regression, and Cox regression analyses performed as part of this study. It could even be argued that given a relatively high degree of intercorrelation between the four facets, an inordinate and unacceptable portion of variance was partialled out of Facets 1, 2, and 3 when Facet 4 was controlled, thereby bringing the results of the incremental validity analyses into question. Although this argument could potentially explain why Facets 1, 2, and 3 had trouble achieving significance in the incremental validity analyses, it fails to explain why these facets also performed poorly in the univariate ROC analyses where partialling played no role. In fact, the ROC institutional aggression analysis for Sample 1, which was the only analysis in which Facets 1 and 2 achieved significance (Facet 3 did not achieve significance

in any of the analyses), was the weakest analysis of the four performed in this study. Conceptually, the institutional aggression analysis for Sample 1 was dominated by verbal aggression, and methodologically, the institutional aggression analysis for Sample 1 was the only analysis that was not fully prospective.

This study is limited in several respects. First, the sample sizes, although more than 100 in all cases, were small in comparison to the 300-plus samples used in the Walters et al. (2008) investigation. Given the relatively low base rates observed for institutional aggression and violent recidivism in this study, several of the analyses may have been lacking in power. In fact, a power analysis of the top performing PCL-R measure in each analysis revealed below threshold (.80) power in three of the four analyses (institutional aggression and violent recidivism in Sample 1 and violent recidivism in Sample 2). Hence, some predictors might have achieved statistical significance had the samples been larger. Second, the outcome measures were derived from official sources (e.g., the FBI's NCIC database). Research indicates that official data are incomplete and potentially misleading (Monahan et al., 2001). Whenever possible, future investigators should obtain outcome data from self-report and collateral contacts as well as official records. Third, institutional adjustment in Sample 1 covered only a 6-month period, and the PCL was administered at the beginning of the first 3-month institutional infraction period. It should be noted, however, that the records review for the PCL was restricted to events occurring prior to the patient's hospital admission, institutional adjustment was restricted to events occurring since the patient's hospital admission, and the PCL and institutional adjustment reviews were conducted by two groups of raters independent of one another. Finally, because the samples for both studies reported in this article comprised forensic psychiatric patients the generalizability of these findings to general prison inmates is uncertain.

Risk assessment measures often cannot be applied across multiple settings, as the conditions that affect the occurrence of institutional misconduct differ in important respects from those in the community (Heilbrun, 2009). Psychopathy has been shown to be related to violent reoffense risk in the community, and to a lesser extent to institutional aggression and violence. The current results may help explain these complex relationships by highlighting certain specific aspects of psychopathy that are most strongly related to violence in these different contexts, if, in fact, Facet 4 and its parcels are a core element of psychopathy. If it is determined that Facet 4 and its parcels are not core elements of psychopathy then we need to know what they are core elements of so that we can develop more efficient research and clinical measures. The current results also remind us of the difficulties associated with predicting low base rate behaviors, the importance of defining aggression

and violence as precisely as possible, and the need to consider non-PCL factors that could potentially augment the ability of the PCL-R to predict violence risk. This will assist future researchers in exploring the implications of particular components of psychopathy as they apply to both prediction and risk management.

Authors' Note

The assertions and opinions contained herein are the private views of the authors and should not be construed as official or as reflecting the views of the Federal Bureau of Prisons or the U.S. Department of Justice.

Declaration of Conflicting Interests

The authors declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding

Notes

1. Exp[B] = exponentiation of the unstandardized regression coefficient, which takes the form of an odds ratio. The numbers in brackets after the value for exp[B] are the lower and upper limits of the 95% confidence interval for exp[B].
2. β = standardized regression coefficient. The number in brackets after the value for β is the standard error of β.

References

Buffington-Vollum, J., Edens, J. F., Johnson, D. W., & Johnson, J. K. (2002). Psychopathy as a predictor of institutional misbehavior among sex offenders: A retrospective replication. *Criminal Justice and Behavior, 29*, 497-511.

Cleckley, H. (1976). *The mask of sanity* (5th ed.). St. Louis, MO: C. V. Mosby. (Original work published 1941)

Conroy, M. A., & Murrie, D. C. (2007). *Forensic assessment of violent risk: A guide for risk assessment and risk management*. Hoboken, NJ: Wiley.

Cooke, D. J., Michie, C., & Skeem. J. (2007). Understanding the structure of the Psychopathy Checklist–Revised. *British Journal of Psychiatry, 190*, s39-s50.

Edens, J. F., Buffington-Vollum, J., Keilen, A., Roskamp, P., & Anthony, C. (2005). Predictors of future dangerousness in capital murder trials: Is it time to "disinvent the wheel?" *Law and Human Behavior, 29*, 55-86.

Edens, J. F., & Douglas, K. S. (2006). Assessment of interpersonal aggression and violence: Introduction to the special issue. *Assessment, 13*, 221-226.

Edens, J. F., Marcus, D. K., Lilienfeld, S. O., & Poythress, N. G. (2006). Psychopathic, not psychopath: Taxometric evidence for the dimensional structure of psychopathy. *Journal of Abnormal Psychology, 115*, 131-144.

Edens, J. F., Petrila, J., & Buffington-Vollum, J. K. (2001). Psychopathy and the death penalty: Can the Psychopathy Checklist–Revised identify offenders who represent a continuing threat to society? *Journal of Psychiatry & Law, 29*, 433-448.

Edens, J. F., Skeem, J. L., & Douglas, K. S. (2006). Incremental validity analyses of the Violence Risk Appraisal Guide and the Psychopathy Checklist: Screening Version in a civil psychiatric sample. *Assessment, 13*, 368-374.

Glover, A. J. J., Nicholson, D. E., Hemmati, T., Bernfeld, G. A., & Quinsey, V. L. (2002). A comparison of predictors of general and violent recidivism among high-risk federal offenders. *Criminal Justice and Behavior, 29*, 235-249.

Grann, M., Långström, N., Tengström, A., & Kullgren, G. (1999). Psychopathy (PCL-R) predicts violent recidivism among criminal offenders with personality disorders in Sweden. *Law and Human Behavior, 23*, 205-217.

Greene, W. H. (2002). *LIMDEP Version 8.0: Econometric modeling guide* (Vol. 2). Plainview, NY: Econometric Software.

Guay, J. P., Ruscio, J., Knight, R. A., & Hare, R. D. (2007). A taxometric analysis of the latent structure of psychopathy: Evidence for dimensionality. *Journal of Abnormal Psychology, 116*, 701-716.

Hare, R. D. (1980). A research scale for the assessment of psychopathy in criminal populations. *Personality and Individual Differences, 1*, 111-119.

Hare, R. D. (1998). The Hare PCL-R: Some issues concerning its use and misuse. *Legal and Criminological Psychology, 3*, 99-119.

Hare, R. D. (2003). *The Hare Psychopathy Checklist–Revised*. Toronto, Ontario, Canada: Multi-Health Systems.

Hare, R. D., & Neumann, C. S. (2006). The PCL-R assessment of psychopathy: Development, structural properties, and new directions. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 58-88). New York: Guilford Press.

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

Heilbrun, K. (1997). Prediction versus management models relevant to risk assessment: The importance of legal decision-making context. *Law and Human Behavior, 21*, 347-359.

Heilbrun, K. (2009). *Evaluation of violence risk in adults*. New York: Oxford University Press.

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.

Kroner, D. G., & Mills, J. F. (2001). The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions. *Criminal Justice and Behavior, 28*, 471-489.

Lynam, D. R., Hoyle, R. H., & Newman, J. P. (2006). The perils of partialling: Cautionary tales from aggression and psychopathy. *Assessment, 13*, 328-341.

Marcus, D. K., Lilienfeld, S. O., Edens, J. F., & Poythress, N. G. (2006). Is antisocial personality disorder continuous or categorical? A taxometric analysis. *Psychological Medicine, 36*, 1571-1582.

Mills, J. F., Kroner, D. G., & Hemmati, T. (2007). The validity of violence risk estimates: An issue of item performance. *Psychological Services, 4*, 1-12.

Monahan, J., Steadman, H., Silver, E., Appelbaum, P., Robbins, P. C., Mulvey, E., et al. (2001). *Rethinking risk assessment: The MacArthur study of mental disorder and violence*. New York: Oxford University Press.

Ruscio, J., Haslam, N., & Ruscio, A. M. (2006). *Introduction to the taxometric method: A practical guide*. Mahwah, NJ: Lawrence Erlbaum.

SPSS. (2008). *SPSS base 17.0 user's guide*. Chicago: Author.

Thomas, L. (1997). Retrospective power analysis. *Conservation Biology, 11*, 276-280.

Walters, G. D. (2003). Predicting institutional adjustment and recidivism with the Psychopathy Checklist factor scores: A meta-analysis. *Law and Human Behavior, 27*, 541-558.

Walters, G. D., Diamond, P. M., Magaletta, P. R., Geyer, M. D., & Duncan, S. A. (2007). Taxometric analysis of the Antisocial Features scale of the Personality Assessment Inventory in federal prison inmates. *Assessment, 14*, 351-360.

Walters, G. D., & Duncan, S. A. (2005). Use of the PCL-R and PAI to predict release outcome in inmates undergoing forensic evaluation. *Journal of Forensic Psychiatry and Psychology, 16*, 459-476.

Walters, G. D., Duncan, S. A., & Geyer, M. D. (2003). Predicting disciplinary adjustment in inmates undergoing forensic evaluation: A direct comparison of the PCL-R and PAI. *Journal of Forensic Psychiatry and Psychology, 14*, 382-393.

Walters, G. D., Gray, N. S., Jackson, R. L., Sewell, K. W., Rogers, R., Taylor, J., et al. (2007). A taxometric analysis of the Psychopathy Checklist: Screening Version (PCL:SV): Further evidence of dimensionality. *Psychological Assessment, 19*, 330-339.

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.

Webster, C. D., Eaves, D., Douglas, K. S., & Wintrup, A. (1995). *The HCR-20 scheme: The assessment of dangerousness and risk*. Vancouver, British Columbia, Canada: Simon Fraser University and British Columbia Forensic Psychiatric Services Commission.

Webster, C. D., Harris, G. T., Rice, M. E., Cormier, C., & Quinsey, V. L. (1994). *The violence prediction scheme: Assessing dangerousness in high risk men*. Toronto, Ontario, Canada: Centre of Criminology, University of Toronto.