DSM-5 Antisocial Personality Disorder: Predictive Validity in a Prison Sample
Online First Publication, September 1, 2014. http://dx.doi.org/10.1037/lhb0000105

CITATION
DSM-5 Antisocial Personality Disorder: Predictive Validity in a Prison Sample

John F. Edens and Shannon E. Kelley
Texas A&M University

Jennifer L. Skeem
University of California–Berkeley

Scott O. Lilienfeld
Emory University

Kevin S. Douglas
Simon Fraser University and Mid-Sweden University

Symptoms of antisocial personality disorder (ASPD), particularly remorselessness, are frequently introduced in legal settings as a risk factor for future violence in prison, despite a paucity of research on the predictive validity of this disorder. We examined whether an ASPD diagnosis or symptom-criteria counts could prospectively predict any form of institutional misconduct, as well as aggressive and violent infractions among newly admitted prisoners. Adult male (n = 298) and female (n = 55) offenders were recruited from 4 prison systems across the United States. At the time of study enrollment, diagnostic information was collected using the Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders (4th ed.; DSM–IV; APA, 1994) Axis II Personality Disorders (SCID-II; First, Gibbon, Spitzer, Williams, & Benjamin, 1997) supplemented by a detailed review of official records. Disciplinary records were obtained from inmates’ respective prisons covering a 1-year period following study enrollment and misconduct was categorized hierarchically as any (general), aggressive (verbal/physical), or violent (physical). Dichotomous ASPD diagnoses and adult symptom-criteria counts did not significantly predict institutional misconduct across our 3 outcome variables, with effect sizes being close to 0 in magnitude. The symptom of remorselessness in particular showed no relation to future misconduct in prison. Childhood symptom counts of conduct disorder demonstrated modest predictive utility. Our results offer essentially no support for the claim that ASPD diagnoses can predict institutional misconduct in prison, regardless of the number of adult symptoms present. In forensic contexts, testimony that an ASPD diagnosis identifies defendants who will pose a serious threat while incarcerated in prison presently lacks any substantial scientific foundation.

Keywords: antisocial personality disorder, risk assessment, prison violence, future dangerousness, capital sentencing

The American Psychiatric Association’s Diagnostic and Statistical Manual of Mental Disorders (DSM) diagnosis of antisocial personality disorder (ASPD) is characterized by an enduring pattern of unlawful behavior, aggressiveness, deceitfulness, impulsivity, irresponsibility, reckless disregard for the welfare of others, and/or remorselessness manifest during adulthood, as well as evidence of conduct disorder (CD) in childhood or adolescence. Despite a radical proposal by the DSM-5 (APA, 2013) Personality Disorder Workgroup to transform the personality disorders into dimensionally informed personality prototypes, the ASPD diagnostic criteria have survived unchanged from DSM–IV (APA, 1994) to DSM-5. Although the DSM-5 (as well as earlier editions) contains cautionary language indicating that psychiatric diagnoses were not developed to meet the needs of the courts or legal professionals, diagnoses of ASPD have been applied to address pragmatically important legal questions.

The ASPD diagnosis is frequently used for forensic purposes in capital murder trials to support the argument that defendants diagnosed with ASPD remain a continuing threat to society and will commit criminal acts of violence in the future if they are not put to death (DeMatteo, Murrie, Anumba, & Keesler, 2011; Edens & Cox, 2012). For example, in one well-known case, a psychiatrist testified that the defendant’s ASPD was “extremely severe” and that he was “one hundred percent certain that he is and will continue to be a threat no matter where he is” (Cook v. State, 1991). Ironically, this defendant ultimately was exonerated and released—after 22 years on death row (Cook, 2008). Such cases are not isolated events: Recent estimates suggest that at least one in every 25 death-sentenced inmates in the United States has been wrongfully convicted (Gross, O’Brien, Hu, & Kennedy, 2014). Although the exact frequency with which diagnoses of ASPD continue to be used by prosecution experts to justify death sentences is unknown, a recent survey of legal professionals has
suggested that it is common and has considerable impact in these cases (Edens & Cox, 2012).

Despite ASPD’s apparent widespread use in the legal system, the reliability and validity of this diagnosis has repeatedly been called into question (Cunningham & Reidy, 1998; Freedman et al., 2013; Morey & Ochoa, 1989). For example, the DSM-5 (APA, 2013) field trials indicated that ASPD had one of the poorest diagnostic interrater reliabilities (Freedman et al., 2013), exacerbating previous concerns that this diagnosis can be unstable across examiners (Cunningham & Reidy, 1998). Poor reliability may stem from the fact that clinicians are prone to inappropriately assign an ASPD diagnosis when patients do not meet sufficient diagnostic criteria, but do have histories of criminal behavior, exploitation of others, vandalism, or school expulsion (Morey & Ochoa, 1989). In forensic and correctional settings, the accuracy of the ASPD diagnosis may also suffer from an excessive emphasis on isolated events (i.e., specific crimes) rather than enduring patterns of behavior (Cunningham & Reidy, 1998). Confidence in ASPD diagnoses is further undermined by evidence of temporal instability, significant criteria overlap with sequelae of substance abuse, and heterogeneous symptom constellations. Related to the latter point, the likelihood of separable subtypes of ASPD (Poythress, Edensm et al., 2010) may attenuate the value of the broader diagnosis in forecasting important outcomes.

Compounding these concerns is a dearth of research examining the predictive validity of ASPD. Although there is evidence that ASPD is frequently comorbid with substance abuse and suicidality (Glenn, Johnson, & Raine, 2013), surprisingly little is known about the diagnosis’s prognostic value for behavioral outcomes such as aggression and violence. Despite the frequency with which this disorder is introduced as evidence of future dangerousness at capital sentencing, there is scant evidence that ASPD diagnoses identify offenders who will engage in serious violence while incarcerated (Edens, Buffington-Vollum, Keilen, Roskamp, & Anthony, 2005). A few recent studies have identified a relationship between categorical and dimensional measures of ASPD and self-reported institutional violence (Warren & Burnett, 2013; Warren et al., 2002) and the disorder has been linked to disciplinary infractions in prison (Toch & Adams, 1986). However, these investigations assessed institutional misconduct retrospectively rather than prospectively and the associations reported may be artifacts of labeling problematic inmates as “antisocial.” That is, interview-based diagnoses of ASPD in these instances may reflect a history of disruptive behavior while incarcerated rather than indicate future risk for aggression or violence.

Indeed, there are strong reasons to doubt a potent association between ASPD and violence. Prevalence rates of ASPD among prisoners are high, typically ranging anywhere between 50%–80%, depending on the individual study. In sharp contrast, base rates of serious violence in prisons are low (Cunningham & Reidy, 2002; Cunningham, Reidy, & Sorensen, 2008; Cunningham & Sorensen, 2007; Edens et al., 2005). Even if ASPD were a significant predictor of aggressive behavior during incarceration, such a combination invariably compromises predictive utility because of an exceedingly high false-positive rate. The majority of individuals with the disorder will not engage in the behavior of interest.

To our knowledge, only one prospective study has examined the relationship between ASPD and prison misconduct. Walters and Knight (2010) assessed newly admitted male inmates for adult and childhood symptoms of ASPD using a structured interview and conducted a brief follow-up (0.1 to 9.5 months) concerning receipt of disciplinary infraction reports while they were incarcerated. Adult antisocial symptoms were associated with greater levels of general institutional misconduct only when two or more symptoms of antecedent CD were also endorsed. In fact, inmates only meeting adult ASPD criteria were indistinguishable from non-ASPD inmates in terms of the likelihood of receiving a disciplinary infraction, suggesting that historical information detailing prior childhood conduct is a salient component in predicting institutional adjustment. Also of note, base rates of violent behavior in this study were too low to perform meaningful analyses—consistent with other research indicating that violent infractions in prison settings are relatively rare phenomena (Cunningham & Reidy, 2002; Cunningham et al., 2008; Cunningham & Sorensen, 2007; Cunningham et al., 2005; Edens et al., 2005).

When attempting to forecast institutional misconduct, focus on ASPD is also problematic in that it may detract attention from known predictors of prison violence (see DeMatteo et al., 2011, Chapter 8), which include such factors as young age, low education, low IQ, prior prison terms, short-term sentences, and prison-gang affiliation (Cunningham & Sorensen, 2006, 2007; Cunningham, Sorensen, & Reidy, 2005; Diamond, Morris, & Barnes, 2012). These findings also suggest that violence risk could be confounded with ASPD diagnoses, given that both are elevated among men, young adults, African Americans, and less educated individuals (Cunningham & Reidy, 1998; Grant et al., 2004; Huebner, 2003). Similarly, a spurious correlation—reflecting an admixture of two distributions—between ASPD and prison violence may be present in samples that include both female and male inmates. ASPD diagnoses are relatively uncommon among women, and women generally commit less (and less severe) violence than men while incarcerated (Harer & Langan, 2001), which creates the appearance of an association between the two variables that is in reality due to gender differences.

To address the limitations of previous research examining the relationship between ASPD and institutional behavior, the present study examined the utility of ASPD diagnoses in prospectively predicting general, aggressive, and violent forms of misconduct among male and female prisoners incarcerated across the United States. Newly admitted inmates were the focus of our investigation, as institutional misconduct tends to occur more frequently within the first few years of confinement, prior to environmental and social adaptation (Flanagan, 1980). Although the main focus of this research is to address the predictive validity of ASPD diagnoses, we also examined three related potential predictors. First, given evidence that ASPD is a dimensional rather than categorical construct (Marcus, Lilienfeld, Edens, & Poythress, 2006), we examined symptom counts to determine whether those who met relatively more of the criteria were more likely to engage in future misconduct than those who met fewer of the criteria. Second, because mental health experts often lack high-quality information about whether examinees met CD criteria during their youth, we examined separately the predictive utility of the adult criteria for the disorder (absent consideration of the CD criterion). We likewise examined the predictive validity of CD symptom counts in isolation, because item-response theory analyses have suggested that the CD criterion for ASPD is most strongly related to the latent antisocial trait tapped by the diagnosis (Jane, Olt-
manns, South, & Turkheimer, 2007). Finally, because (a) expert testimony concerning a defendant’s alleged remorselessness is frequently provided to support an argument that he or she is particularly dangerous (Weisman, 2007), and (b) remorselessness is a specific symptom within the ASPD criteria set, we examined its independent utility in predicting prison misconduct.

Method

Participants

Participants were 353 male (n = 298) and female (n = 55) incarcerated offenders who were interviewed using the Structured Clinical Interview for DSM–IV Axis II Personality Disorders (SCID-II; First, Gibbon, Spitzer, Williams, & Benjamin, 1997) as part of a larger study funded by the United States National Institute of Mental Health (NIMH) that examined antisocial and psychopathic personality traits among prisoners and substance abusers. Detailed information regarding this larger sample and additional measures administered has been described elsewhere (Poethress, Edens, et al., 2010; Poethress, Lilienfeld, et al., 2010). Participants were general-population inmates recruited from multiple medium- and maximum-security prison sites in Florida, Nevada, Oregon, and Utah. Recruitment targeted newly admitted inmates for whom official records of institutional misconduct were obtained after a follow-up period of 1 year. The mean age of the sample was 30.43 years (SD = 6.82). Most participants self-identified as Caucasian (56.7%; 41.1% African American; 2.2% not indicated; 6.5% of participants additionally self-identified as ethnically Hispanic).

Measures

SCID-II. The ASPD module of the SCID-II (First et al., 1997) is a semistructured interview that assesses adult and childhood DSM–IV (APA, 1994) criteria for ASPD. This measure is a widely used diagnostic instrument and exhibits adequate levels of inter-rater reliability for total symptom count and categorical ratings of ASPD (Maffei et al., 1997). Interrater reliability for this research project was established through observation of randomly selected SCID-II interviews conducted by research assistants, in addition to requisite file reviews. All observations were performed by the project manager (KSD), whose SCID-II scoring was treated as the criterion against which research assistant scoring was compared. The project manager traveled to each site approximately every 6 months and observed two cases per visit, for a maximum of six visits. Concordance for this project was good for ASPD diagnoses ($\kappa = .74, n = 50$) and total symptom count (intraclass correlations [ICC] = .86, n = 46).

Research assistants scored SCID-II items dichotomously as either present or absent, based on information extracted during a clinical interview and from a detailed review of participants’ institutional files. A slight majority of participants in this sample met criteria for ASPD (52.10%). Consistent with previous literature, the prevalence was much higher among men (56.9%) than women (25.5%). Rates of the ASPD diagnosis also varied across race, with more African Americans (61.1%) meeting criteria than Caucasians (45.5%). The average number of adult symptoms met was 3.67, with considerable variability evident within our sample (SD = 1.97, range = 0 to 7).

Institutional misconduct. Disciplinary records were obtained from participants’ respective prisons covering a period of 1 year after study enrollment. Disciplinary infraction policies were obtained from each state’s department of correction and compared to develop a standard coding system for categorizing documented misconduct. To be consistent with prior meta-analytic research on institutional infractions (e.g., Guy, Edens, Anthony, & Douglas, 2005), we coded infractions into three hierarchical outcome variables: (a) general infractions of any type (e.g., possession of contraband), (b) aggressive infractions (including both physical and verbal aggression), and (c) physically violent infractions (e.g., assault or battery with a deadly weapon). The prevalence of misconduct over the course of the study was relatively low, with the mean (SD) number of infractions being 1.33 (3.33) for general, .73 (2.54) for aggressive, and .05 (.27) for physically violent, and the modal number being 0 for each category. As such, the outcome measures were dichotomized (0 = no infractions; 1 = one or more infractions). Over the 1-year follow-up period, 40.2% of our sample incurred at least one general infraction and 25.5% incurred at least one aggressive infraction. The base rate for physically violent infractions was low (4.2%), which is in keeping with other research in this area (e.g., Walters & Knight, 2010). As such, the aggressive infraction category primarily represents acts of verbal aggression, though not exclusively.

Procedure

Participants at each site were recruited randomly from lists of general-population inmates meeting basic inclusion criteria, which required that participants speak English and self-identify as African American or Caucasian. We elected not to recruit individuals who self-identified primarily as members of other racial groups, given the paucity of evidence regarding the validity of ASPD and psychopathy in these groups. In addition, individuals currently receiving psychotropic medication for active symptoms of psychosis and those with an estimated IQ screening of less than 70 (Ammons & Ammons, 1962) were excluded. The study protocol was administered by research assistants who were advanced clinical psychology or social work graduate students and who were extensively trained on data-collection procedures by the senior investigators. After a complete description of the study, written informed consent was obtained using procedures approved by university institutional review boards. Approximately 85% of individuals approached to participate in the larger study chose to do so. The protocol took on average 4.5 hours to complete (typically over two or three sessions). Participants received $20 as compensation for their time.

Analytic Plan

In recent years, the area under the curve (AUC) derived from receiver-operating-characteristic (ROC) curve analyses has gained ascendency as the effect size of choice in risk-assessment research (Mossman, 1994). Although AUC values are informative regarding the performance of an instrument or predictive tool in general, they usually do not directly address the question of how accurate predictions of a criterion are when they are provided, because AUC values typically are not derived from any one particular cut-off score. Instead, they provide a global indicator of the per-
formance of the predictor variable across all possible values of that predictor variable. Although not commonly reported, AUCs can be computed for individual cut-off scores, in which case they represent the midpoint between sensitivity and specificity for that particular score.

A more important concern about AUC values is that, when attempting to predict low base-rate events, high values can be obtained even when predictions of a criterion occurring are wrong most of the time, due to the disproportionate number of false positives evident among those predicted to engage in the outcome of interest (e.g., future violent behavior; Streiner, 2003). As such, a consideration of classical test-theory statistics (e.g., sensitivity, specificity), particularly those that are influenced by base rates (e.g., positive predictive power [PPP], negative predictive power [NPP]), is also necessary when judging the accuracy of predictions of misconduct among prisoners. In particular, rates of PPP quantify how frequently predictions of future violence are correct when based on the presence of an ASPD diagnosis.

**Results**

Table 1 summarizes the performance of ASPD diagnoses as predictors of the three types of disciplinary infractions. Across all three types of misconduct, diagnoses did not significantly predict these outcomes, with global performance indicators (phi, kappa) being close to zero in absolute magnitude. The differential rates of PPP and NPP reported across the criterion measures are chiefly a function of the differing base rates of the three criterion measures under investigation. Although ROC curves most commonly are used to examine predictive validity across an entire range of test scores or diagnostic symptom counts, we computed these values for the ASPD diagnostic cut-off point and obtained similarly negligible results: AUC ($SE$) = .55 (any = .03), .52 (aggressive = .04); .51 (violent = .08).

Table 1 provides results based on diagnostic status independent of the presence/absence of the CD criterion for ASPD. Disregarding the need to meet this criterion resulted in a higher prevalence of ASPD diagnoses (68.6%), which resulted in some minor differences in terms of classification statistics. Such differences are essentially moot, however, given that none of the global statistics was statistically significant. ROC results were similar to the negligible phi and kappa results reported in Table 1: AUC ($SE$) = .52 (.03, any); .53 (.04, aggressive); and .49 (.08, violent).

Next, we addressed whether predictive validity might be improved by examining symptom counts rather than dichotomous ASPD diagnoses. For the adult-criteria symptom count, biserial correlations with the three outcome measures were $r = .16, .04$ (aggressive), and .01 (violent), all $p$ values nonsignificant. AUC ($SE$) values from ROC analyses were .55 (.03) for any, .52 (.03) for aggressive, and .52 (.07) for violent infractions, again all nonsignificant. We next examined the childhood symptom counts of CD. Unlike the adult criteria, these demonstrated modest evidence of predictive utility, with correlations of $r = .16, .12, p < .05$ (aggressive), .11, $p < .05$ (violent). Corresponding AUC ($SE$) values were: .60 (.03), $p < .01$ (any); .59 (.04), $p < .05$ (aggressive); .67 (.06), $p < .05$ (violent). We then examined the utility of a total symptom count that combined both CD and adult criteria into a single variable. This resulted in effect sizes that were virtually identical to (or slightly lower than) those achieved by the childhood symptom count in isolation.

Next, we examined the lack of remorse criterion in relation to each outcome. Phi coefficients ranged between $-0.04$ and .00, as did kappa values (all $p$ values nonsignificant). Similarly, AUC values were all nonsignificant, ranging from .48 to .51.

Because our sample included a relatively small percentage of female inmates, we re-conducted the analyses after dropping women from our sample. Among male inmates, results were virtually identical to those reported for the full sample. Other than weak trends for childhood-criteria symptom counts to relate to misconduct ($r$ values between .10 and .14; AUC values between .57 and .68), ASPD was not a significant predictor of misconduct across our three outcome measures.

Because dichotomizing our criterion measures could have resulted in a loss of statistical power, we also conducted the preceding set of analyses using both negative binomial and Poisson regression methods, which are appropriate when outcome measures are count data. Results of these analyses were highly similar to those reported for the dichotomous outcome measures. Further details regarding these findings are available from the first author.

**Supplemental Analyses**

Given the overall weak results and lack of utility of the remorselessness criterion, we conducted exploratory analyses on the remaining adult criteria to investigate whether individual symptoms might be of some use in forecasting institutional misconduct.

<p>| Table 1 | Diagnostic Efficiency Statistics for ASPD Diagnoses in Predicting Institutional Misconduct |
|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|</p>
<table>
<thead>
<tr>
<th>Statistic</th>
<th>Any</th>
<th>Aggressive</th>
<th>Violent</th>
<th>Any</th>
<th>Aggressive</th>
<th>Violent</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPP</td>
<td>44.6</td>
<td>27.2</td>
<td>4.3</td>
<td>41.7</td>
<td>26.9</td>
<td>4.1</td>
</tr>
<tr>
<td>NPP</td>
<td>64.5</td>
<td>76.3</td>
<td>95.9</td>
<td>63.1</td>
<td>77.5</td>
<td>95.5</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>57.7</td>
<td>55.6</td>
<td>53.3</td>
<td>71.1</td>
<td>72.2</td>
<td>66.7</td>
</tr>
<tr>
<td>Specificity</td>
<td>51.7</td>
<td>49.0</td>
<td>47.9</td>
<td>33.2</td>
<td>32.7</td>
<td>31.4</td>
</tr>
<tr>
<td>Overall</td>
<td>45.9</td>
<td>49.3</td>
<td>51.8</td>
<td>51.6</td>
<td>57.2</td>
<td>67.1</td>
</tr>
<tr>
<td>Phi</td>
<td>.09</td>
<td>.04</td>
<td>.01</td>
<td>.05</td>
<td>.05</td>
<td>−.01</td>
</tr>
<tr>
<td>Kappa</td>
<td>.09</td>
<td>.03</td>
<td>.00</td>
<td>.04</td>
<td>.03</td>
<td>.00</td>
</tr>
</tbody>
</table>

*Note.* PPP = positive predictive power; NPP = negative predictive power; Overall = overall correct classification rate. None of the inferential statistical values reported is significant at $p < .05$. 
only individual criterion demonstrating any utility was “failure to conform to social norms with respect to lawful behaviors,” which yielded significant phi values for any and aggressive infractions (.18 and .16, respectively, both p values < .01), but only marginally significant AUC values (.56, SE = .03, ps < .10). The base rate for this criterion in our sample was .88, however, which, not surprisingly, resulted in PPP values of only .44 for any infractions and .28 for aggressive infractions. In essence, in the rare cases in which this criterion was not present, its absence was a good indicator that an infraction would not occur (corresponding NPP values were .84 and .93), but the presence of this criterion was an accurate predictor of misconduct only 44% of the time for any infraction and only 28% of the time for an aggressive infraction.

Finally, given the generally poor performance of ASPD diagnostic information to predict misconduct, one might question whether the criterion measures themselves might be problematic operationalizations of institutional adjustment problems. To investigate this issue, we examined the association between our outcome measures and correlates of institutional adjustment problems identified in earlier large-scale studies and meta-analyses (Cunningham & Sorensen, 2007; Cunningham et al., 2005; Diamond et al., 2012). These variables were age, education level, IQ screening scores, gender, and race. Although we do not wish to suggest that any of these variables should be used for risk-assessment purposes in practice, they are demonstrated correlates of disciplinary infractions in prisons and thereby provide an opportunity to examine whether our criterion measures display a similar pattern of associations with other well-established predictor variables. As can be seen in Table 2, several significant, modest effect-size estimates were obtained with the three indicators of misconduct, indicating that our criterion measures possess sufficient enough reliability and validity to be statistically predicted by established risk factors.

We also examined these predictor variables using both negative binomial and Poisson regression, which again demonstrated results very similar to when the dichotomous criterion measures were examined. In fact, effects were somewhat stronger than those reported in Table 2 for the dichotomized criterion measures. Further detail is available from the first author upon request.

Discussion

The use of psychiatric diagnoses to influence the disposition of legal cases has a long history. In particular, diagnoses related to antisocial or psychopathic traits have been used to justify punitive sanctions (e.g., indeterminate commitment, capital punishment) for offenders, based in part on arguments that these diagnoses foretell violence and other criminal behavior (Cunningham & Reidy, 1998; DeMatteo et al., 2011; Sevilla, 1999). We sought to address whether an ASPD diagnosis identified offenders newly admitted to prison who would engage in acts of institutional misconduct during their first year of incarceration. Such findings can inform the justice system about the utility of this diagnosis in identifying high-risk inmates, which bears on sentencing decisions and security classifications.

Our results offer essentially no support for the claim that ASPD diagnoses provide meaningful information concerning the institutional maladjustment of prisoners. The adult criteria were not useful for identifying who would commit disciplinary infractions over our 1-year follow-up. Remorselessness in particular showed no relevance to identifying those who would engage in problematic behavior over this time period, despite being commonly cited as an important risk factor for future aggression (Cunningham & Reidy, 2002).

Similar to the recent findings of Walters and Knight (2010), the only aspect of ASPD diagnoses that showed any potential utility for forecasting misconduct was CD symptom count. This finding is consistent with research demonstrating an association between more severe childhood antisocial conduct and risk of criminal and violent behavior as an adult (Farrington, 1991; Mason et al., 2004; Robins, 1978), even among generally at-risk groups (e.g., patients with schizophrenia; Swanson et al., 2006). In one study of emotional and behavioral problems, childhood conduct problems (e.g., attacking people, destroying things) as reported by 10-year-old children and their parents and teachers best predicted a pattern of serious violence in early adulthood (Mason et al., 2004). Further, specific constellations of adult antisocial behaviors may correspond to the number of behavioral disturbances evidenced in childhood (Robins, 1978).

The overall poor performance of ASPD diagnoses in our study does not bode well regarding their predictive utility in U.S. prisons. These findings are consistent with previous meta-analytic work demonstrating low predictive validity for measures of psychopathy within U.S. prison settings (Guy et al., 2005), although effect sizes were somewhat stronger in non-U.S. prisons and forensic facilities. One could argue that our study provided an optimal opportunity to demonstrate a relationship between ASPD diagnoses and prison misconduct, if such a relationship were to exist. Our research assistants conducted extensive interviews and had access to institutional file data concerning participants. Our reliability data indicated that the SCID-II ASPD module (First, Gibbon, Spitzer, Williams, & Benjamin, 1997) yielded consistent scores across examiners—much more so than what has been demonstrated for this diagnosis in the DSM-5 field trials (APA, 2013; Freedman et al., 2013). Moreover, our criterion measures were based on comprehensive institutional records of misconduct and displayed expected patterns of association with

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of misconduct</td>
</tr>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Education level</td>
</tr>
<tr>
<td>IQ</td>
</tr>
</tbody>
</table>

| Phi/Kappa (SE) |
| Any | Aggressive | Violent |
| Gender | .08/.05 (.03) | .11/.05 (.02)** | -.03/.00 (.01) |
| Race | .13/.04 (.01)* | .15/.04 (.01)** | .15/.02 (.01)** |

Note. Due to some missing data points, sample sizes for these analyses range from 341 to 353. Education level was coded as 1 = no high school diploma, 2 = completed GED, 3 = high school diploma, 4 = some college completed, 5 = completed college. Gender was coded 1 = female, 2 = male. Race was coded 1 = Caucasian, 2 = African American. *p < .05. **p < .01.
known predictors of prison misconduct (e.g., age, IQ) that were generally of a magnitude similar to what has been reported in other large-scale studies.

If ASPD diagnoses are not useful indicators of risk in prison settings, what factors should be considered? As noted earlier, numerous individual risk factors have been demonstrated to be relevant to institutional risk, although some (e.g., race/ethnicity) raise important ethical/legal concerns in regard to their potential use in legal decision-making. Some individual-level risk factors (e.g., age, education, prior prison confinement) have been incorporated into actuarial systems that have been shown in large-scale studies to classify prisoners into meaningful subgroups in terms of risk level (Cunningham & Sorensen, 2006). Further, several structured professional judgment (SPJ) measures have been shown to predict violence within prison and forensic settings (for reviews, see Otto & Douglas, 2010). The key point is that most robust risk assessment approaches do not rely on any single construct, but attempt to incorporate an array of risk factors.

Although not the focus of this particular study, another important domain widely known to influence rates of prison violence is the role of the situational and environmental factors, which almost surely constraints the rates of physical aggression behind bars (Gadon, Johnstone, & Cooke, 2006; Sorensen, Cunningham, Vig, & Woods, 2011; Tewksbury, Connor, & Denney, 2014). Forensic practitioners working in this area might consider the PRISM model (promoting risk interventions by situational management; Cooke & Johnstone, 2010) as a means of moving beyond simple individual-level risk factors when they engage in the process of institutional risk assessment and management. Although an exhaustive review goes beyond the scope of this article, in brief, the PRISM model focuses attention on empirically supported situational/environmental factors, such as structural aspects (e.g., supervision and security level), staff features (e.g., age, experience), temporal aspects (e.g., timing of assaults), location, crowding, organizational management practices and relationships (e.g., relationships between administrators and frontline staff), and program availability for inmates. Models such as PRISM might lead to more nuanced risk-assessment and management strategies that go beyond a primary focus on the individual offender.

Although our study benefitted from several methodological strengths, there were also important limitations we should note that may inform future work in this area. First, the rates of physically violent infractions were low, raising concerns about the stability of our results for this criterion measure. That said, our study is typical in this regard, in that the incidence of serious acts of violence in prison tends to be low, even among inmates who have committed violent crimes (e.g., capital murder) in the community (Cunningham & Reidy, 2002; Cunningham et al., 2008; Cunningham & Sorensen, 2007; Cunningham et al., 2005; Edens et al., 2005). Second, our sample did not contain a large sample of female offenders, which constrains the generalizability of our results. Third, it is unclear to what extent our results would generalize to less secure contexts, especially if the base rate of the “failure to conform to social norms with respect to lawful behaviors” symptom is not so close to 100%, as it was in our data. Fourth, given Guy et al.’s (2005) findings of lower predictive validity for the closely related concept of psychopathy within U.S. prisons than non-U.S. prisons, our findings should not be considered to generalize beyond the U.S. Fifth, our follow-up period was only one year in length. Although results might have differed over a more extended time frame, such an outcome seems unlikely given that (a) we captured the (early) period of maximum risk for prisoners to engage in misconduct and (b) risk decreases considerably the longer prisoners are incarcerated. Finally, one could argue that our means of operationalizing the severity of ASPD (i.e., total symptom counts) does not correspond to what forensic practitioners mean when they describe a defendant as having a “severe case” of this disorder. Although there may be some merit to this criticism, the burden would fall on such practitioners to provide evidence that their alternative means of quantifying severity has predictive utility.

In conclusion, the results of this study offer little if no support for the claim that ASPD diagnoses identify inmates who go on to engage in institutional misconduct in prison. In forensic contexts, the use of this diagnosis to support claims that defendants will be a serious threat to others, even if incarcerated in prison, would seem to rest on dubious empirical grounds.

References


Received May 30, 2014
Revision received July 24, 2014
Accepted July 24, 2014

View publication status