Published in final edited form as:

Psychol Assess. 2011 March ; 23(1): 174–183. doi:10.1037/a0021290.

Predictive and Incremental Validity of the Violence Risk Appraisal Guide Scores with Male and Female Jail Inmates

Mark E. Hastings,

Department of Psychology, George Mason University, Loudoun County Mental Health Center,
Private Practice

Shilpa Krishnan,

Department of Psychology, George Mason University

June P. Tangney, and

Department of Psychology, George Mason University

Jeffrey Stuewig

Department of Psychology, George Mason University

Abstract

The present study examines the predictive and incremental validity of Violence Risk Appraisal Guide scores in a sample of 328 male and 145 female jail inmates held on felony charges. Significant gender differences were observed in VRAG item and total score means, as well as in correlations between the VRAG and concurrent measures of aggression. VRAG scores significantly predicted institutional misconduct during incarceration and recidivism in the first year post-release for male inmates, but not for female inmates. In terms of incremental validity, VRAG scores predicted institutional misconduct and recidivism beyond that accounted for by psychopathy for male inmates, but not for female inmates. Implications for clinical practice and future research are discussed.

Keywords

Violence Risk Appraisal Guide; validity; inmates; gender; recidivism

The assessment of risk for future violence has become a staple of forensic psychological practice (Monahan, 1996). The advent of actuarial risk instruments such as the Violence Risk Appraisal Guide (VRAG; Quinsey, Harris, Rice, & Cormier, 1998) have contributed mightily to this endeavor. The VRAG is a 12-item actuarial risk assessment tool that has been validated for use in a wide variety of populations such as sex offenders (Harris, Rice, Quinsey, Lalumiere, Boer, & Lang, 2003; Langton, Barbaree, Seto, Peacock, Harkins, & Hansen, 2007), civil psychiatric patients (Harris, Rice & Camilleri, 2004), mentally disordered (Gray, Fitzgerald & Taylor, 2007) and non-North American offender samples (Doyle, Dolan, & McGovern, 2002; Urbaniok, Noll, Grunewald, Steinbach & Endrass,

Correspondence concerning this article may be addressed to Mark E. Hastings, Department of Psychology, George Mason University, 4400 University Drive, Fairfax, VA 22030.

Publisher's Disclaimer: The following manuscript is the final accepted manuscript. It has not been subjected to the final copyediting, fact-checking, and proofreading required for formal publication. It is not the definitive, publisher-authenticated version. The American Psychological Association and its Council of Editors disclaim any responsibility or liabilities for errors or omissions of this manuscript version, any version derived from this manuscript by NIH, or other third parties. The published version is available at www.apa.org/pubs/journals/PAS

2006; Kroner, Stadtland & Eidt, 2007). The ability of the VRAG to predict violent behavior among criminal and mentally-disordered male inmates has been well-established (Glover, Nicholson, Bernfeld, & Quinsey, 2002; Kroner & Mills, 2001).

Females represent one of the fastest growing population of inmates (Bureau of Justice Statistics, 2006), yet the field of violence risk assessment has not kept pace with this particular sub-group of inmates. There are no actuarial risk assessment instruments designed specifically for use with female inmates, and the validity of interpretation of VRAG scores with female inmates remains largely unknown. Female inmates share some risk factors for violence and recidivism in common with male inmates, specifically psychopathy (Loucks & Zamble, 1999; Salekin, Rogers, Ustad, & Sewell, 1998) and prior criminal history (Bonta, Pang, & Wallace-Capretta, 1995). However, research also indicates that female inmates are not simply “male inmates in drag,” but rather possess unique treatment needs and pathways to violence and offending (Bonta, Pang, & Wallace-Capretta, 1995). For example, Van Voorhis, Wright, Salisbury, and Bauman (2010) found that gender-responsive variables make for stronger predictive schemes among female inmates compared to gender-neutral variables alone. They also found that long-established predictors of recidivism among male inmates, such as criminal thinking and antisocial associates, were much less relevant to recidivism among female inmates. Yet instruments, such as the VRAG, are often used in correctional settings despite the fact that they were not originally created, or been extensively validated, for females.

Walters (2006) conducted a meta-analysis of the predictive utility of actuarial risk assessment instruments versus self-report measures. Six samples included in the analysis were female or male/female mixed samples. Overall, the mean weighted effect size for male versus female studies was comparable. However, there was greater variability in the range of effect sizes in the female sample studies, and none of the female samples utilized the VRAG. Campell, French and Gendreau's (2009) more recent meta-analysis examined the predictive utility of a broader range of risk instruments and other psychological measures, including the VRAG. The VRAG performed well in predicting institutional violence and violent recidivism. However, of the effect sizes where gender composition of the samples was reported, the vast majority were generated from male samples. It was not possible to examine gender as a moderator of predictive utility for any of the risk assessments, including the VRAG.

Only three studies to date have specifically assessed the predictive validity of VRAG scores in female samples. Harris, Rice, and Cormier (2002) reported on the predictive accuracy of the VRAG in a small sample of female forensic patients embedded within a larger mixed gender sample. They found the VRAG was unrelated to subsequent violent behavior in the 59 female patients in the sample. However, they also noted the base rate of violence in the female sample was low. Examination of the individual VRAG items for females revealed that the victim injury (item 8) was inversely related to subsequent violence and the Hare Psychopathy Checklist-Revised (PCL-R) (item 12) was unrelated. Information on the other VRAG items was not reported and as such were probably not significantly related to subsequent violent behavior for females.

Harris, Rice, and Camilleri (2004) examined VRAG data on 423 male and 318 female civil psychiatric patients who participated in the MacArthur Violence Risk Assessment Project. A modified VRAG was employed in this study owing to a lack of information needed to score several VRAG items. It was found that the abbreviated VRAG predicted violence as well for women (AUC = .73) as for men (AUC = .71), with similar rates of violence evidenced by men and women (27% and 23%, respectively). However, when Edens, Skeem, and Douglas (2006) conducted incremental validity analyses on this same data set, they found that for the

sample as a whole, modified VRAG scores contributed minimally to the prediction of violence once the PCL:SV was controlled. No gender-specific incremental validity analyses were conducted.

Coid et al. (2009) examined the predictive validity of scores from five risk assessment instruments with both male and female offenders, including the VRAG. The VRAG was predictive of violent (AUC = .65), general (AUC = .66), and any (AUC = .66) recidivism in female offenders; however, the level of predictive accuracy was small to moderate and incremental validity analyses were not conducted. Lastly, the VRAG was significantly less predictive of violent recidivism for female versus male offenders.

Only one study to date has examined the utility of the VRAG with a female correctional sample, and no study has examined how the individual VRAG items may differ in their relationship, both concurrently and prospectively, with relevant constructs depending on gender. Also understudied is the incremental validity of VRAG scores above and beyond Hare psychopathy scores separately by gender.

The present study attempts to fill these gaps in the literature by examining the concurrent, predictive, and incremental validity of the VRAG total score as well as the individual items. We directly compare the performance of these items for males versus females. Understanding the differential relationships, if any, may help improve our use of risk appraisal procedures for women.

METHOD

Participants

Participants were 473 adults (328 males and 145 females) recently incarcerated at a county jail just outside Washington, DC, who were held on at least one felony charge or serving a sentence of at least four months. All participants were housed in medium or maximum security general population. Participants completed the Personality Assessment Inventory (PAI; Morey, 1991, 2007) using “touch-screen” computers that require minimal familiarity with computers (e.g., no keyboard, no mouse). In addition to presenting questionnaire items visually, the computer read each item aloud to all participants via headphones, thus accommodating participants with limited reading proficiency.

Table 1 shows the demographics of the sample. Participants ranged in age from 18.2 to 69.7, and there was a significant gender difference with female inmates being on average 3.5 years older than male inmates. Education ranged from 5 to 19 years, and there was a modest but statistically significant gender difference with female inmates reporting on average nearly one more year of education than male inmates. Participant IQ (Wonderlic, 1999) ranged from 64 to 138, with no difference between men and women. PCL:SV psychopathy scores ranged from 1 to 24. Consistent with much previous research, men scored higher than women. Regarding race, both male and female samples were diverse, but there was no gender difference.

Measures and Procedures

Shortly after assignment to the general population, eligible inmates were presented with a description of the study and asked to participate. Inmates were assured of the voluntary and confidential nature of the project. In particular, it was emphasized that the decision to participate or not would have no bearing on their status at the jail nor their release date. Of those approached, 74% agreed to participate and provided written consent. The main reason cited for non-participation was anticipated imminent release. There was no gender difference in rate of consent. Of those who consented, 82% ($n = 518$) remained in the jail

long enough to complete portions of the assessment relevant to this report. The main reason for non-completion was being bonded (bailed) out of jail. Of the 518, 5% ($n = 29$) were excluded because of validity concerns based primarily on two validity scales from the PAI – Inconsistency and Infrequency. Specifically, participants were dropped if either Inconsistency or Infrequency was above the recommended cut-offs (T scores of 72 and 74, respectively, Morey, 1991) and the other was considered elevated (above 69). Second, interviewers routinely reported if there were validity concerns during data collection. In cases where interviewer concerns were raised, data were further examined for response bias, response sets, elevation of other validity scales, and documentation of problems from other sessions. Finally, 16 participants (all males) were Spanish only speakers who completed the PAI in face to face interviews instead of on the touch screen computer. Due to the different mode of administration, plus the gender difference in the group, these individuals were subsequently dropped from these analyses – leaving a final sample for analysis of 473. Regarding confidentiality, interviews were conducted in the privacy of professional visiting rooms, used by attorneys, or secure classrooms. In addition, data were protected by a Certificate of Confidentiality from DHHS. Inmates who agreed to participate, and who completed the 4–6 session intake assessment, received a \$15–18 honorarium.

Hare Psychopathy Checklist—Screening Version (PCL:SV; Hart, Cox, & Hare, 1995) was used to assess psychopathy. Completion of this 12-item checklist requires an in-depth psychosocial history interview, videotaped for coding with the participant's permission. In addition, as part of participation, inmates granted access to criminal and jail records. This collateral information was used by trained clinicians in conjunction with the interview to code the PCL-SV.

Prior to coding the PCL:SV, interviewers completed an advanced graduate course on theory, research, and assessment of psychopathy, including intensive supervised training in the administration and scoring of the PCL-R and PCL:SV. Only those who successfully met interrater reliability criteria for both forms were cleared for coding study protocols. A randomly selected set of 54 cases were double-coded by a referent clinician (the first author) who brought to the project 15 years of professional experience conducting forensic psychological evaluations, as well as advanced training and experience in the administration and scoring of the PCL-R and PCL:SV. Single measure intraclass correlations, using a one-way random effects model, were .85, .79, and .85 for Part 1, Part 2, and Total PCL:SV scores, respectively, showing a high degree of interrater reliability.

Violence Risk Appraisal Guide—(VRAG; Harris, Rice, & Quinsey, 1993; Quinsey et al., 1998). The VRAG is a 12-item actuarial risk assessment instrument developed using a sample of mentally disordered offenders. Similar to the PCL:SV, we assessed interrater reliability with a single measure intraclass correlation using a one-way random effects model. The randomly selected 52 cases showed high reliability ($ICC = .89$).

Personality Assessment Inventory—(PAI; Morey, 1991, 2007). The following scales from the PAI were utilized in concurrent validity analyses: the Antisocial (ANT) and Aggression (AGG) scales, the Aggression-Physical (AGG-P) subscale, and the Violence Potential Index (VPI). The ANT and AGG subscales have been repeatedly found to be associated with institutional misconduct among male and female criminal offenders (Edens & Ruiz, 2006; Skopp et al., 2007) and with recidivism (Boccaccini et al., 2010; Walters, 2006). Although less thoroughly researched, the VPI has also shown promise in predicting one-year post-release recidivism (Hastings et al., 2003).

Institutional misconduct was coded using jail records obtained between time of enrollment in the study and release date ($N=315$ males, $N=123$ females). Three indices of institutional

misconduct were calculated: 1) *Number of Incidents* recorded in the inmate's file regardless of whether they led to a formal institutional charge or finding; 2) *Number of Formal Charges* levied against the inmate for violation of institutional rules; and 3) *Number of Formal Physical Charges* for physical violence (e.g., fighting, assault on a correctional officer, etc.) levied against the inmate. For certain analyses variables were dichotomized to reflect any incident or charge. In the current sample, 173 (40%) inmates had at least one incident (127 males and 46 females, $\chi^2 = .32, ns$), 137 (31%) inmates had formal charges against them (104 males and 33 females, $\chi^2 = 1.6, ns$), and 32 (7%) inmates had physical charges against them, with males being significantly more likely to have a physical incident than females (29 males, 3 females, $\chi^2 = 5.98, p < .05$). In another analysis, we considered the Number of Incidents of institutional misconduct ($M = 1.38, SD = 2.77$ for the total sample). More incidents of institutional misconduct were recorded for male inmates ($M = 1.54, SD = 3.02$) compared to female inmates ($M = .97, SD = 1.98, t(335.2) = -2.31, p < .05$). On average participants were incarcerated for 151 days ($SD = 108$, range 3 to 536). Men ($M = 160$ days, $SD = 114$) were incarcerated for a significantly longer period of time than women ($M = 124$ days, $SD = 88, t(285.3) = -3.59, p < .001$). Once length of incarceration was taken into account there was no gender difference in number of institutional misconduct incidents $F(1, 435) = .73, p = .39$.

One-Year Post-Release Recidivism—At one-year post-release, 290 participants (207 males and 83 females) were assessed, providing self-reports of arrests, and self-reports of criminal behavior for which they were not arrested (violent and non-violent). Interviewers were blind to participants' prior PCL:SV psychopathy scores and VRAG scores. Analyses indicated no baseline differences between those who were followed up versus those who were not in gender, race, age, education, IQ, VRAG, Psychopathy, Antisocial (ANT), Aggression (AGG), Aggression-Physical (AGG-P), and the Violence Potential Index (VPI).

Four indices of self-reported recidivism were calculated: 1) *Number of Arrests* reflects the number of different types of crimes for which the participant was arrested (a measure of criminal versatility); 2) *Number of Undetected Offenses* reflects the number of different types of criminal acts participants engaged in that were not detected by authorities (a measure of criminal versatility); 3) *Number of Arrests or Undetected Offenses* is a sum score of the previous two variables; and 4) *Number of Violent Arrests or Undetected Offenses* reflects the number of different types of violent crimes for which the participant was arrested plus the number of different types of undetected violent criminal acts. Violent arrests and acts were defined as assault, domestic violence, robbery, murder, and kidnapping, which matches up closely with the original outcome violence variable of interest used in the development of the VRAG (all assaultive behaviors, sex offenses, armed robbery, forcible confinement, and threatening violence with a weapon). For certain analyses variables were dichotomized to reflect any arrest or offense.

The recidivism rate at one year post-release was 64%, with significantly more males ($n = 146, 70.5\%$) than females ($n = 41, 49.4\%$) reporting either being arrested or having committed an undetected crime ($\chi^2 = 11.55, p < .01$). The rate of violent recidivism was 17%, with significantly more males ($n = 42, 20\%$) than females ($n = 8, 9.6\%$) reporting having been arrested or having committed undetected violent criminal acts ($\chi^2 = 4.78, p < .05$).

RESULTS

Gender Differences in VRAG Item and Total Score Means

We first examined whether male and female inmates differed in their scores on the VRAG. The overall VRAG Total score was significantly higher for male versus female inmates (see Table 2). Of the 12 individual VRAG items, significant gender differences were observed on

seven items. (Chi-Squares are presented for dichotomous items; Mann Whitney tests are presented for ordinal items, e.g., Item 2 and 3). Specifically, at the time of the index offense women tended to be older (Item 7) and more likely to have been married (Item 4). They had less severe histories of elementary school problems (Item 2) and alcohol problems (Item 3), lower levels of psychopathy (Item 12), less severe injury to victims (Item 8), and the victim in their index offense was slightly less likely to be female (Item 9). We also examined gender differences on the total of the 11 VRAG items that essentially augment the Hare Psychopathy scores. Results indicated that these “Hare-less” VRAG scores for male inmates ($M=8.05$, $SD=6.64$) were significantly higher than female inmates ($M=5.61$, $SD=5.84$, $t(471)=-3.83$, $p < .01$).

Gender Differences in Correlations between the VRAG and Other Measures of Aggression

Concurrent correlations between VRAG Total score and other measures of aggression or violence potential were substantially and significantly correlated for male inmates (See Table 3). Concurrent correlations between VRAG Total score and other measures of aggression and violence potential were significantly correlated for female inmates as well; however, the magnitude of the correlations was smaller. Tests of independent correlations confirmed that the VRAG to Antisocial, VRAG to Physical Aggression, and the VRAG to the Violence Potential Index correlations were significantly higher for male compared to female inmates ($p < .05$).

To further investigate the source of these gender differences in the correlates of the VRAG total score, we examined the relationship of each VRAG item to the antisocial, aggression, physical aggression, and violence potential scales from the PAI separately for males and females. As shown in Table 3, the VRAG items were less predictive for females compared to males. Only two VRAG items were significantly correlated with all concurrent measures of aggression for both genders: Item 2 History of Elementary School Problems and Item 12 Hare psychopathy. Six VRAG items showed consistently higher convergent validity with other measures of aggression for men compared to women: Item 3 (history of alcohol problems), Item 4 (never married), Item 5 (non-violent criminal history), Item 6 (failure on conditional release), Item 7 (age at index offense), and Item 9 (no female victim). Tests of independent correlations, however, did not show many significant differences between the correlations for males versus females on the individual items, except for Item 9 (see Table 3). The differential validity of Item 9 as a function of gender is especially notable; a high score on Item 9 (no female victim) was consistently positively correlated with measures of antisocial and aggressive behavior among men, but for women, this item was consistently *negatively* related to other measures of antisocial behavior and aggression. Two VRAG items showed poor validity for both men and women. Item 8 (Victim Injury) and Item 11 (No DSM-III-R Diagnosis of Schizophrenia) were largely unrelated to concurrent measures of aggression, likely because of the minimal variance observed in these items.

Correlations and ROCS with Institutional Misconduct

Among male inmates, the VRAG Total score was significantly correlated with Number of Incidents ($r=.28$, $p<.01$), Number of Formal Charges ($r=.27$, $p<.01$), and Number of Formal Physical Charges ($r=.17$, $p<.01$), whereas among female inmates, the VRAG Total score was not significantly correlated with any measure of institutional misconduct (r 's $=.05$, $.01$, $-.04$). The correlation between VRAG Total score and Number of Incidents was significantly different between male and female inmates ($Z=-2.48$, $p<.05$), as was the correlation between VRAG Total score and Number of Formal Charges ($Z=-2.6$, $p<.05$), as well as for Number of Formal Physical Charges ($Z=-2.0$, $p<.05$).

For males, Receiver Operating Characteristics (ROC) revealed moderate predictive accuracy for the VRAG Total score with each of the three indices of institutional misconduct: Any Incident, Any Formal Charge, and Any Formal Physical Charge (see Table 4). In contrast, ROC analyses indicated that the VRAG Total score was not significantly predictive of institutional misconduct among female inmates.

To further explore these gender differences in predictive validity, correlations between each individual VRAG item and the three indices of institutional misconduct were conducted (See Supplemental Table 1). Item 1 (Lived with Biological Parents till age 16), Item 2 (Elementary School Problems), Item 5 (Nonviolent criminal history), and Item 12 (Hare Psychopathy Score) were significantly correlated with institutional misconduct for male inmates. Magnitudes ranged from small to moderate. In contrast, none of the individual VRAG items were significantly correlated with any of the three indices of institutional misconduct for female inmates. Tests of independent correlations show that males and females only significantly differed on Number of Incidents for Items 1 and 2, and Number of Formal Charges for Items 1 and 12 (see Supplemental Table 1).

Correlations and ROCs with One-Year Post-Release Recidivism

The VRAG Total score was significantly moderately correlated with Number of Arrests ($r=.25, p<.01$), Number of Undetected Offenses ($r=.38, p<.01$), Number of Arrests or Undetected Offenses ($r=.39, p<.01$), and Number of Violent Arrests or Undetected Offenses ($r=.37, p<.01$) for male inmates, whereas VRAG Total score had lower correlations with the measures of recidivism for the female inmates ($r=.21, p=.054$; $r=.27, p<.05$; $r=.27, p<.05$; $r=.17, p=.13$). These differences between female and male inmates were not statistically significant ($Z=-.32, p=.75$, $Z=-.93, p=.35$, $Z=-1.02, p=.31$, $Z=-1.64, p=.10$, respectively).

For males, Receiver Operating Characteristics (ROC) revealed moderate predictive accuracy for the VRAG Total score and all four indices of one-year recidivism (Table 4). In contrast, VRAG Total score was only significantly related to Any Arrest or Undetected Offense for females.

Correlations between each individual VRAG item and the four indices of one-year recidivism were conducted (See Supplemental Table 2). No item was consistently significantly related to all the one-year post-release recidivism variables for male inmates. However, Items 2 (Elementary School Problems), 3 (History of Alcohol Problems), 5 (Nonviolent criminal history), 6 (Failure on Conditional Release), and 12 (Hare Psychopathy Score) showed some consistent pattern of correlations. In general, for male inmates the magnitudes ranged from small to moderate. In contrast, only Item 12 (Hare Psychopathy Score) and Item 2 (Elementary School Problems) were significantly correlated with any of the four indices of one-year post-release recidivism for female inmates. However, no correlations between female and male inmates were statistically significant from each other.

Bin Base Rates

The base rate of violent reoffending in the VRAG normative sample was 31% after seven years. The current study base rate for violent reoffending after one year was 17%. The initial VRAG validation sample was divided into nine bins according to range of scores. Figure 1 shows the observed recidivism rates associated with each VRAG bin from the current male and female samples alongside the recidivism rates from the male VRAG validation sample. Despite the significantly shorter follow-up period and a limited number of subjects in bins 1–3 and 8–9, the overall trajectory of the recidivism rates for male inmates was consistent with that of the VRAG normative sample. That is, for VRAG bins 4–9, increasingly greater recidivism rates were observed for each bin. However, the stability of the observed

recidivism rates for bins 8–9 is limited given the small number of subjects in each bin. For female inmates, scores were clustered around bins 4–6 and were associated with a small increase in recidivism across the three bins. None of the six females in bin 7, however, evidenced any violent recidivism. A Mann-Whitney test confirmed a significant relationship between the bins and violent recidivism for males ($Z=-5.1, p < .01$) but not for females ($Z=-.78, p=.43$).

The current study base rate for physical institutional misconduct was only 7%. Figure 2 shows the observed rate of physical institutional misconduct for each VRAG bin for the current sample. The low base rate of physical institutional misconduct in the sample, particularly for females where only three incurred physical institutional misconduct charges, rendered analysis of the VRAG at the bin level relatively uninformative.

Incremental Validity of VRAG Scores

To examine the incremental validity of VRAG scores we used the PCL:SV score to predict the modified version of the VRAG total score (i.e., all items summed except for the psychopathy score item) and saved the residuals. This residual score is the independent variance of the modified VRAG once the variance associated with the PCL:SV score has been removed. Similar to Edens, et al., (2006) we used this residual score in a series of ROC analyses to examine what the VRAG contributes to institutional misconduct and post-release recidivism above and beyond the PCL:SV.

As can be seen in Table 5, the modified VRAG scores continued to predict outcomes even after controlling for psychopathy for men (although the relationship was reduced compared to full VRAG in Table 4). For women, however, the modified VRAG scores were unrelated to outcomes. In addition to examining the dichotomous variables, we ran correlations of the modified VRAG residual with the continuous variables of institutional misconduct and post-release recidivism. For males, the residual score positively and significantly related to institutional misconduct (mean $r = .14$; range .13 – .15) and to recidivism (mean $r = .27$; range .21 – .33) showing incremental validity. For females, the residual score was not significantly related to institutional misconduct (mean $r = -.04$; range -.03 to -.04) or recidivism (mean $r = .09$; range .07 – .09).

DISCUSSION

Consistent with previous research, the VRAG proved to be a moderately effective predictor of several indices of institutional misconduct and one-year post-release recidivism among male inmates. Most important, the VRAG – an actuarial measure designed to predict risk for violence – was a significant predictor of physical institutional charges as well as violent recidivism for males, but not for females. While not specifically validated for non-violent offenses, the VRAG also significantly predicted general institutional misconduct and detected and undetected general post-release recidivism for men, but again not for women.

The gender differences in predictive validity can be explained in part by examining the individual VRAG items and their relationship to concurrent measures of antisociality and aggression. In the current study, ten of the twelve VRAG items were significantly correlated with concurrent measures of antisociality, aggression, and violence potential in male inmates. However, only two of the individual VRAG items (i.e., Item 2: Elementary School Problems; Item 12: Hare Psychopathy Score) reached statistical significance in female inmates. Most of the non-significant individual VRAG item correlations for female inmates were in the same direction but lesser magnitude as for male inmates, but two items demonstrated either no relationship or an inverse relationship to concurrent measures of aggression in female inmates. Specifically, Item 9: No Female Victim was inversely related

to aggression, and Item 4: Marital Status was not related to aggression in female inmates. Female to female violence may have a completely different meaning than male to female violence thus accounting for the lack of concurrent validity for Item 9. And while a good predictor of violence and crime for males, the influence of marital status on females' behavior may be more dependent on the type of partner she is in a relationship with than the actual fact of being married. For example, Moffit et al. (2001) found that as antisocial girls aged they were more likely to become involved with antisocial men, which in turn reinforced and increased their own antisocial behavior. Research has also found that marriage serves as a protective factor against suicidality in males, but not in females (Masocco et al., 2008), and being unmarried predicted lower life span in males, but only to a lesser degree in females (Kaplan & Kronick, 2006).

Two VRAG items were poor predictors for both male and female inmates. Item 8: Victim Injury was largely unrelated to concurrent measures of aggression, and Item 11 where those who did not have a diagnosis of Schizophrenia were often lower on aggression. Because seriously psychotic inmates were excluded from general population there were almost no inmates in this sample with a diagnosis of schizophrenia which may have contributed to the contrary finding. It is possible, however, that this would be true in most general population settings and so the utility of this item may be sample specific. Nonetheless, it appears that the lack of predictive validity of VRAG scores in female inmates stems at least in part from the lack of relationship between the individual VRAG items and measures of aggression.

One limitation of the current study is that some of the outcome variables demonstrated low base rates of occurrence. However, many of the outcome base rates were low for both genders, yet the VRAG showed a medium effect size in predicting aggression among male offenders. The results of this study may also be limited to female correctional samples, as opposed to, for example, civil or forensic psychiatric samples that may vary in demographic composition and/or in base rates for some items. Finally, the reports of recidivism were all self-report and as such there may be a bias toward false negatives in the current sample lowering validity estimates. Although official records also have many limitations of their own, it would be best to have information from multiple sources so as to guard against method bias.

The literature has largely been silent concerning the use of the VRAG with female populations, yet given a lack of actuarial risk assessment instruments normed specifically on female populations, it is not unusual for some individual forensic experts and agencies to advocate the use of the VRAG with females given the proven fallibility of unguided clinical prediction (Grove & Meehl, 1996; Meehl, 1954). The findings of this study clearly indicate such use is inadvisable. At best, the VRAG Total score provides clinicians with no more useful information about the violence risk of female inmates beyond that already accounted for by psychopathy and at worst, the VRAG provides a distorted, less accurate assessment of females' violence risk, given its reliance on individual items that bear little predictive relevance to females' antisociality and aggression.

In short, these findings argue strongly *against* the use of the VRAG for assessing violence risk among female offenders. The current study demonstrates that while some violence risk factors are gender-neutral (i.e., criminal history), clinicians cannot simply assume that many of the risk factors for violence (or risk assessment instruments) developed on male inmates will be relevant to risk assessment in female populations. One potential alternative is structured professional judgment tools such as the HCR-20 and LSI-R, which have demonstrated some reliability and validity in female samples (Folsom & Atkinson, 2007; Nicholls et al., 2004). Future research should also attempt to identify female-specific valid

indicators of risk for violence that can augment or supplement the predictive utility of psychopathy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors wish to thank all of the lab members at the Human Emotions Research Laboratory at George Mason University and the inmates who participated in this project. James E. Hastings, Ph.D. is also thanked for his valuable comments on earlier drafts of this manuscript. The authors also wish to acknowledge the National Institute of Drug Abuse for funding the research described in this manuscript. All opinions expressed in this article are the sole opinions of the authors and do not reflect opinions or policy of George Mason University, the Loudoun County Mental Health Center, or any other agency.

References

- Boccaccini MT, Murrie DC, Hawes SW, Simpler A, Johnson J. Predicting recidivism with the Personality Assessment Inventory in a sample of sex offenders screened for civil commitment as sexually violent predators. *Psychological Assessment*. 2010; 22(1):142–148. [PubMed: 20230160]
- Bonta J, Pang B, Wallace-Capretta S. Predictors of recidivism among incarcerated female offenders. *The Prison Journal*. 1995; 73:277–294.
- Campbell MA, French S, Gendreau P. The prediction of violence in adult offenders: A meta-analytic comparison of instruments and methods of assessment. *Criminal Justice and Behavior*. 2009; 36:567–590.
- Coid, Yang, Ullrich, Zhang, Sizmur, Roberts, Farrington, Rogers. Gender differences in structured risk assessment. *Journal of Consulting and Clinical Psychology*. 2009; 77(2):337–348. [PubMed: 19309193]
- Doyle M, Dolan M, McGovern J. The validity of North American risk assessment tools in predicting in-patient violent behaviour in England. *Legal and Criminological Psychology*. 2002; 7(2):141–154.
- Edens JF, Ruiz MA. On the validity of validity scales: The importance of defensive responding in the prediction of institutional misconduct. *Psychological Assessment*. 2006; 18(2):220–224. [PubMed: 16768599]
- Edens JF, Skeem JL, Douglas KS. Incremental validity analyses of the Violence Risk Appraisal Guide and the Psychopathy Checklist: Screening Version in a civil psychiatric sample. *Assessment*. 2006; 13(3):368–374. [PubMed: 16880286]
- Folsom J, Atkinson JL. The generalizability of the LSI-R and the CAT to the prediction of recidivism in female offenders. *Criminal Justice and Behavior*. 2007; 34(8):1044–1056.
- Glover AJJ, Nicholson DE, Bernfeld GA, Quinsey VL. A comparison of predictors of general and violent recidivism among high-risk federal offenders. *Criminal Justice and Behavior*. 2002; 29(3): 235–249.
- Gray NS, Fitzgerald S, Taylor J. Predicting future reconviction in offenders with intellectual disabilities: The predictive efficacy of VRAG, PCL-SV, and the HCR-20. *Psychological Assessment*. 2007; 19(4):474–479. [PubMed: 18085940]
- Grove WM, Meehl PE. Comparative Efficiency of Informal (Subjective, Impressionistic) and Formal (Mechanical, Algorithmic) Prediction Procedures: The Clinical-Statistical Controversy. *Psychology, Public Policy, and Law*. 1996; 2(2):293–323.
- Hare, RD. *The Hare Psychopathy Checklist - Revised*. Toronto: Multi-Health Systems; 1991.
- Harris GT, Rice ME, Camilleri JA. Applying a forensic actuarial assessment (the Violence Risk Appraisal Guide) to nonforensic patients. *Journal of Interpersonal Violence*. 2004; 19(9):1063–1074. [PubMed: 15296617]
- Harris GT, Rice ME, Cormier CA. Prospective replication of the Violence Risk Appraisal Guide in predicting violent recidivism among forensic patients. *Law and Human Behavior*. 2002; (4):377–394. [PubMed: 12182529]

- Harris GT, Rice ME, Quinsey VL. Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behavior*. 1993; 20:315–335.
- Harris GT, Rice ME, Quinsey VL, Lalumiere ML, Boer D, Lang C. A multisite comparison of actuarial risk instruments for sex offenders. *Psychological Assessment*. 2003; 15(3):413–425. [PubMed: 14593842]
- Hart, SD.; Cox, DN.; Hare, RD. The Hare Psychopathy Checklist: Screening Version. Toronto: Multi-Health Systems; 1995.
- Hastings, ME.; Stuewig, J.; Drapalski, A.; Tangney, JP. Psychopathy, Violence Risk Assessment, and the Personality Assessment Inventory (PAI). Paper presented at the annual meeting of the American Society of Criminology; Denver, CO. 2003 November.
- Kaplan RM, Kronick RG. Marital status and longevity in the United States population. *Journal of Epidemiology and Community Health*. 2006; 60(9):760–765. [PubMed: 16905719]
- Kroner DG, Mills JF. The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions. *Criminal Justice and Behavior*. 2001; 28(4):471–489.
- Kroner C, Stadtland C, Eidt M. The validity of the Violence Risk Appraisal Guide (VRAG) in predicting criminal recidivism. *Criminal Behaviour and Mental Health*. 2007; 17(2):89–100. [PubMed: 17295202]
- Langton CM, Barbaree HE, Seto MC, Peacock EJ, Harkins L, Hansen KT. Actuarial assessment of risk for reoffense among adult sex offenders: Evaluating the predictive accuracy of the Static-2002 and five other instruments. *Criminal Justice and Behavior*. 2007; 34(1):37–59.
- Loucks A, Zamble E. Predictors of recidivism in serious female offenders. *Corrections Today*. 1999; 61:26–32.
- Masocco M, Pompili M, Vichi M, Vanacore N, Lester D, Tatarelli R. Suicide and marital status in Italy. *Psychiatric Quarterly*. 2008; 79(4):275–285. [PubMed: 18600458]
- Meehl, PE. Clinical versus statistical prediction: A theoretical analysis and a review of the evidence. Minneapolis: University of Minnesota Press; 1954.
- Moffitt, T.; Caspi, A.; Rutter, M.; Silva, PA. Sex, antisocial behavior, and mating: Mate selection and early child bearing. In: Moffitt, T.; Caspi, A.; Rutter, M.; Silva, PA., editors. Sex differences in antisocial behavior: Conduct Disorder, delinquency, and violence in the Dunedin longitudinal study. New York: Cambridge University Press; 2001. p. 184–197.
- Monahan J. Violence prediction: The past twenty and the next twenty years. *Criminal Justice and Behavior*. 1996; 23(1):107–120.
- Morey, LC. Personality Assessment Inventory: Professional manual. Odessa, FL: Psychological Assessment Resources; 1991.
- Morey, LC. Personality Assessment Inventory: Professional manual. 2. Odessa, FL: Psychological Assessment Resources; 2007.
- Nicholls TL, Ogloff JRP, Douglas KS. Assessing violence risk among male and female civil psychiatric patients: The HCR-20, PCL:SV, and VSC. *Behavioral Sciences and the Law*. 2004; 22:127–158. [PubMed: 14963884]
- Quinsey, VL.; Harris, GT.; Rice, ME.; Cormier, CA. Violent offenders: Appraising and managing risk. Washington, DC: American Psychological Association; 1998.
- Salekin RT, Rogers R, Ustad KL, Sewell KW. *Law and Human Behavior*. 1998; 22(1):109–128. [PubMed: 9487793]
- Skopp NA, Edens J, Ruiz MA. Risk factors for institutional misconduct among incarcerated women: An examination of the criterion-related validity of the Personality Assessment Inventory. *Journal of Personality Assessment*. 2007; 88(1):106–117. [PubMed: 17266421]
- Urbanik F, Noll T, Grunewald S, Steinbach J, Endrass J. Prediction of violent and sexual offences: A replication study of the VRAG in Switzerland. *Journal of Forensic Psychiatry & Psychology*. 2006; 17(1):23–31.
- Van Voorhis, Wright, Salisbury, Bauman. Women's risk factors and their contribution to existing risk/need assessment: The current status of a gender-responsive supplement. *Criminal Justice & Behavior*. 2010; 37:261–288.

- Walters GD. Risk-appraisal versus self-report in the prediction of criminal justice outcomes: A meta-analysis. *Criminal Justice and Behavior*. 2006; 33(3):279–304.
- Wonderlic, Inc. Wonderlic Personnel Test & Scholastic Level Exam User's Manual. Libertyville, Ill: Wonderlic Inc; 1999.

Figure 1.
Observed 1-year post-release recidivism from current sample compared to the 7-year recidivism data from original VRAG sample

Figure 2.
Observed physical institutional misconduct rates from current sample

Table 1

Demographic Characteristics and Criminal History of Jail Inmates

Demographic	Total		Males		Females		<i>t-test</i>
	Mean (SD)		Mean (SD)		Mean (SD)		
Age	32.8 (10.0)		31.7 (9.8)		35.2 (10.2)		3.5**
Yrs of Education	11.9 (2.1)		11.7 (2.1)		12.4 (2.1)		3.7**
IQ	94.3 (13.7)		94.0 (13.5)		94.8 (14.3)		.59
Have Psychopathy	12.3 (4.8)		13.2 (4.8)		10.4 (4.3)		-6.2**
Demographic	Total		Males		Females		χ^2
	N (%)		N (%)		N (%)		
Race							5.7
African-American	217 (46)		155 (47)		62 (43)		
Caucasian	178 (38)		113 (35)		65 (45)		
Hispanic	26 (5)		21 (6)		5 (3)		
Other/Mixed	52 (11)		39 (12)		13 (9)		

*
p < .05;**
p < .01

Table 2

Gender differences in individual VRAG items

	Male	Female	
Item 1: Not lived with both parents till 16	68.3%	74.5%	$\chi^2 = 1.8$
Item 2: Elementary school problems			$Z = -5.18^{**}$
None	28.7%	51.7%	
Slight	37.5%	32.4%	
Severe	33.8%	15.9%	
Item 3: History of alcohol problems			$Z = -2.99^{**}$
None	28.7%	34.5%	
1–2	32.0%	44.1%	
3	19.5%	11.0%	
4–5	19.8%	10.3%	
Item 4: Marital status (never married)	34.1%	15.2%	$\chi^2 = 17.8^{**}$
Item 5: Nonviolent criminal history			$Z = -.28$
0	12.2%	13.1%	
1–2	10.1%	7.6%	
3+	77.7%	79.3%	
Item 6: Failure on conditional release	72.3%	71.0%	$\chi^2 = .07$
Item 7: Age at index offense			$Z = -2.93^{**}$
39+	24.1%	40.0%	
34–38	14.9%	12.4%	
28–33	15.9%	10.3%	
27	2.4%	4.1%	
≤ 26	42.7%	33.1%	
Item 8: Victim injury			$Z = -2.06^*$
Death	.6%	0.0%	
Hospitalized	2.1%	0.0%	
Treated/Released	2.7%	1.4%	
None/Slight	94.5%	98.6%	
Item 9: No female victim	84.5%	91.7%	$\chi^2 = 4.6^*$
Item 10: DSM-III-R Personality disorder	4.0%	4.1%	$\chi^2 = .01$
Item 11: DSM-III-R No Schizophrenia	97.9%	98.6%	$\chi^2 = .25$
Item 12: Hare psychopathy score			$Z = -5.66^{**}$
1–4	5.2%	8.3%	
5–7	6.7%	17.9%	
8–10	17.4%	25.5%	
11–16	43.9%	40.7%	
17–22	26.8%	6.9%	
23–24	0.0%	0.7%	
VRAG Total Score	8.5 (7.9)	4.8 (6.7)	$t(471) = -4.9^{**}$

Note. $N=328$ (males); $N=145$ (females);

*
 $p < .05$;

**
 $p < .01$.

Chi-squares are presented for dichotomous items, Mann Whitney tests for ordinal items and t-test for continuous items

Table 3
Concurrent Correlations Between VRAG Total & Item Score Means and PAI scales

	PAI-ANT		PAI-AGG		PAI-AGG-P		PAI VPI Index	
	M	F	M	F	M	F	M	F
Item 1	.17**	.06	.21**	.19*	.19**	.16	.18**	.05
Item 2	.33**	.22**	.37**	.33**	.39**	.30**	.29**	.26**
Item 3	.29**	.11	.31**	.12	.33**	.12	.34**	.20*
Item 4	.19**	.01	.09	.02	.14**	.03	.14**	-.08
Item 5	.24**	.14	.16**	.12	.17**	.10	.25**	.11
Item 6	.21**	.15	.17**	.16	.16**	.14	.24**	.15
Item 7	.17**	.10	.20**	.10	.25**	.14	.11*	.01
Item 8	.01	-.08	-.05	.01	-.03	-.02	-.05	.01
Item 9	.14**	-.20*	.07	-.14	.08	-.13	.04	-.09
Item 10	.17**	.21*	.14*	.11	.19**	.11	.14*	.16
Item 11	-.07	-.03	-.10	-.03	-.10	.00	-.15*	-.02
Item 12	.44**	.39**	.37**	.24**	.35**	.19*	.42**	.35**
VRAG Total	.54**	.37**	.50**	.39**	.53**	.36**	.49**	.32**

Note. N=325-326 (males); N=143 (females);

*
p <.05;

**
p <.01.

Correlations that differ significant by gender are presented in bold.

Table 4

VRAG ROC's with Institutional Misconduct and One-Year Post-Release Recidivism

Institutional Misconduct		
	Male Inmates (n=315)	Female Inmates (n=123)
	AUC (95% CI)	AUC (95% CI)
Any Incidents	.65** (.59–.71)	.56 (.45–.66)
Any Formal Charges	.70** (.64–.76)	.53 (.41–.64)
Any Physical Charges	.67** (.59–.75)	.43 (.04–.81)
Recidivism		
	Male Inmates (n=205–207)	Female Inmates (n=83)
	AUC (95% CI)	AUC (95% CI)
Any Arrest	.68** (.60–.75)	.62 (.49–.75)
Any Undetected Offense	.67** (.59–.74)	.61 (.49–.73)
Any Arrest/Undetected Offense	.70** (.62–.78)	.66* (.54–.78)
Any Violent Arrest/Undetected Offense	.76** (.68–.85)	.66 (.47–.85)

Table 5

ROC's of Modified VRAG Controlling for PCL:SV in Order to Examine Incremental Validity

Institutional Misconduct		
	Male Inmates (n=315)	Female Inmates (n=123)
	AUC (95% CI)	AUC (95% CI)
Any Incidents	.62** (.55-.68)	.50 (.40-.61)
Any Formal Charges	.64** (.57-.70)	.53 (.41-.64)
Any Physical Charges	.65** (.57-.74)	.48 (.09-.86)
Recidivism		
	Male Inmates (n=205-207)	Female Inmates (n=83)
	AUC (95% CI)	AUC (95% CI)
Any Arrest	.63** (.56-.71)	.52 (.39-.65)
Any Undetected Offense	.60* (.52-.68)	.50 (.38-.63)
Any Arrest/Undetected Offense	.60* (.52-.69)	.53 (.40-.65)
Any Violent Arrest/Undetected Offense	.71** (.62-.80)	.55 (.40-.71)