

Using the Violence Risk Appraisal Guide (VRAG) to Predict In-Prison Aggressive Behavior in a Swiss Offender Population

Jérôme Endrass

Astrid Rossegger

Andreas Frischknecht

Thomas Noll

Frank Urbaniok

Criminal Justice System, Psychiatric/Psychological Service, Zurich

The present study is a first-time evaluation of the Violence Risk Appraisal Guide's (VRAG) predictive quality for institutional violence in a German-speaking country. The VRAG was assessed for 106 violent and sexual offenders based on their files. Violent infractions during imprisonment were evaluated using the files of the state penitentiary. Results show in accordance with previous studies only a moderate effect between VRAG scores and institutional misconduct. However, these findings were only significant for participants with a sex crime as index offense. In the study, the VRAG was unable to predict verbal and physical violence by violent offenders. The implications of these findings for institutional risk management and the future development of intramural detection of participants at risk in the German-speaking part of Europe are discussed.

Keywords: *violent infractions; VRAG; prediction; offenders; prison*

Introduction

Both in Anglo-Saxon as well as German-speaking countries the scientific debate about whether reoffending is validly predictable is ongoing (Andrews & Bonta, 2003; Horstkotte, 1986; Jockusch, 1990; Kühl & Schumann, 1989; Rasch, 1994; Urbaniok, Noll, Grunewald, Steinbach, & Endrass, 2006).

Authors' Note: Please address correspondence to Thomas Noll, MD, JD, Criminal Justice System, Psychiatric/Psychological Service, P.O. Box Feldstr. 42, 8090 Zurich, Switzerland; e-mail: thomas.noll@ji.zh.ch.

The principal methods of assessing the risk for violent reoffending are the intuitive method, the statistical method (actuarial risk assessments), and the clinical method (mainly using structured clinical guides). In Canada, England, and the United States, the actuarial method predominates.

Even though many studies have been able to demonstrate the superior predictive accuracy of both actuarial methods (Bonta, Law, & Hanson, 1998; Grove, Zald, Lebow, Snitz, & Nelson, 2000; Hilton & Simmons, 2001; Mossman, 1994; Rice & Harris, 1997; Swets, Dawas, & Monahan, 2000) and systematic content-relevant self-report procedures (Walters, 2006) over clinical and intuitive methods, most mental health professionals in Germany, Switzerland, and Austria still use these methods (Urbanik et al., 2006).

Because violent and aggressive behavior in prison is considered to be a constant problem in most penal institutions, there is an ongoing debate concerning the usefulness of actuarial instruments in predicting intramural violence. Acts of violence or aggression not only pose a threat to prison staff and fellow inmates but are also a source of additional cost. An early identification of inmates who tend toward violent or aggressive behavior in prison should therefore be a central goal to both protect the staff and other inmates from becoming victims of violent actions and to lower the overall costs of the penal system.

Once the inmates at risk of behaving violently have been identified, appropriate actions can be taken to reduce the negative effect of institutional misconduct.

The Violence Risk Appraisal Guide (VRAG), developed by Harris, Rice, and Quinsey (1993), is potentially a useful tool for assessing the risk of aggressive behavior in prison. The VRAG was originally designed to predict violent recidivism and was statistically derived from items collected from a psychiatric sample of 618 male offenders detained in Canadian maximum security hospitals. By means of multivariable analyses, those items that had the strongest predictive power were identified. As it turned out, the VRAG was able to predict extramural violent recidivism moderately well for follow-up periods of up to 10 years (Grann, 1998; Rice & Harris, 1995, 1997). The area under the curve (AUC) of the receiver operator characteristics (ROC) was around 0.75 (Harris, Rice, & Cormier, 2002; Rice & Harris, 1995). Replication studies confirmed the reliability and validity for nonpsychiatric incarcerated males (Loza & Dhaliwal, 1997). Grann (1998) reported a ROC of 0.68 for a Swedish prison sample, and Urbanik et al. (2006) found a ROC of 0.73 when they applied the VRAG to a Swiss offender sample.

All the aforementioned research investigated exclusively the risk of extramural recidivism. Thus far, few studies have examined the usefulness of the VRAG for predicting institutional misconduct. Doyle, Dolan, and McGovern (2002) found a ROC of 0.71 for the VRAG to predict inpatient violence for inmates of a medium security unit in England. Nichols, Vincent, Whitemore, and Ogloff (1999) found a significant correlation between VRAG scores and inpatient aggressive behavior within the first 3 months of hospitalization for a sample of forensic psychiatric patients.

Kroner and Mills (2001) compared the VRAG to other risk assessment instruments and found it slightly more accurate for predicting institutional misconduct than the other instruments.

McBride (1999) found the VRAG to be significantly superior compared to the Psychopathy Checklist–Revised (PCL-R) in predicting institutional misconduct and subsequent violence among federally sentenced offenders.

Edens, Skeem, and Douglas (2006) on the other hand found the VRAG scores to be only modestly correlated with violent acts among inmates of a psychiatric facility. Furthermore, they argued that the validity of the VRAG can mainly be attributed to one single item—the PCL-R score.

Even though only a few studies have addressed the subject, there is some empirical evidence to support the value of the VRAG in predicting violence in prison. No investigation has evaluated the predictive validity for intramural violence in the German language area, even though the VRAG proved to predict (extramural) recidivism well in this language area (Urbaniok et al., 2006).

The purpose of the present study was to examine the accuracy of the VRAG for predicting intramural violent and aggressive acts in a Swiss prisoner population.

Method

Participants

We included all male offenders in the largest penitentiary in Switzerland (the Zurich State Penitentiary, Pöschwies) who were sentenced to at least 10 months incarceration and who were administrated by the Zurich correctional and probation service in August 2000.

Further inclusion criteria were the conviction for a violent or sexual offense in the Canton of Zurich (an urban and prealpine area with a population of approximately 1,200,000) and that a psychiatric expert opinion had been given. In all, 123 sexual and violent offenders fulfilled these inclusion criteria. VRAG scores were assessed for the remaining participants as long as the type of offense allowed for it. Excluded were those participants sentenced for committing arson or a noncontact sexual offense and those who had more than four items that could not be coded due to lack of information. The final sample consisted of 106 participants.

Measure and Procedure

The VRAG is comprised of 12 items, including one PCL-R score. Each item is weighted, with the PCL-R item having the greatest range (from -5 to $+12$). The total score is the sum of the 12 items and ranges from -26 to $+40$. According to Rice, Harris, and Cormier (1992), interrater correlation for the items are greater than .80, whereas kappas for the items are greater than .70 (Rice, Harris, & Quinsey, 1990).

In the present study, if the number of missing items was less than five, the final score was corrected using the algorithm described by Quinsey et al. (1998); if the number of missing items was equal to or greater than five, it was excluded from the sample.

The coding of the VRAG and the assessment of a series of psychiatric, psychological, criminological, and socioeconomic variables were based on files and expert opinions. In fact, all data were collected solely from documentary material, without any direct contact with the offenders. Direct contact with the offenders is not necessary and can even decrease predictive accuracy (Quinsey, Harris, Rice, & Cormier, 1998). These files provided the raters with extensive historical details on the individual, such as criminal history, exact type and circumstances of the violent offense, and Axis I and II diagnoses (if any) of the participant.

For information on institutional misconduct, the files of the state penitentiary were consulted. The operationalized definition of violent infractions was intramural physical behavior that harmed or had the potential to harm staff members or fellow inmates. Verbally aggressive behavior (e.g., threats) was coded separately, and we did not include damage to property in our definition. Illegal drug abuse, the possession of illegal drugs, and the total number of disciplinary infractions on the other hand were considered in our analyses.

Four psychologists who had undergone intensive training in the instruments in question collected the data. VRAG scores were assessed for each participant based on data obtained from files. Prior to data assessment, interrater reliability was computed in a pilot study using Krippendorff's alpha (Krippendorff, 1980), a measure that takes chance into account and allows the calculation of reliability coefficients for different scales of measurement as well as for more than two observers. An alpha must have a value between 0.60 and 0.80 to be moderately reliable and between 0.80 and 1.00 to be highly reliable. The results yielded an alpha of 0.89 for the VRAG total score for a sample of $n = 10$ observations.

Statistical Analysis

To determine the predictive validity of the VRAG, violent infractions were estimated as a function of the VRAG in a negative binomial regression model. The natural log of the number of days spent in prison was included as a covariate in the model to equate participants for time at risk. All models were estimated using STATA SE 9.1.

Results

Sociodemographic and Offense-Related Variables

All participants were male, 58% ($n = 66$) were Swiss citizens and 13% ($n = 14$) were from an EU country. At the time of offense, 14% ($n = 16$) were married, 37%

($n = 43$) were in a long-term relationship, and 34% ($n = 39$) had at least one child. In addition, 30% ($n = 35$) lived in a foster home before the age of 15, and 10% ($n = 11$) had been sexually abused during childhood. Nearly 83% had previously been convicted ($n = 96$), and 30% ($n = 35$) had previously been treated in an inpatient psychiatric facility.

The mean age at the beginning of the sentence was 36.1 years ($SD = 8.9$, range = 20 to 60). More than two thirds of the index offenses were violent acts: In nearly half of the sample (48%, $n = 56$) the index offense was murder, attempted murder, manslaughter, or attempted manslaughter. There were 7 cases (6%) of physical assault, 12 (10%) armed robberies, and 5 cases (4%) of arson. In the group of sex offenders, there were 16 cases of rape (14%) and 5 (4.3%) of coercion. In 10% ($n = 11$) of these, the offender abused a child. Finally, 4 cases (4%) who committed violent offenses matched none of these categories.

Disciplinary Infractions

The mean time spent in the institution was 1,693 days ($SD = 1,085.15$, range = 20 to 5,166), 85% ($n = 98$) had at least one disciplinary infraction during incarceration, and the mean number of infractions was 5.3 ($SD = 5.86$, range = 0 to 26). Of the offenders, 15% ($n = 17$) were reported for cannabis use or possession and 13% ($n = 15$) for use or possession of other drugs. In 26% ($n = 30$) of cases a verbally aggressive infraction was listed in the files ($M = 0.54$, $SD = 1.42$, range = 0 to 12), and about one third of the participants (30%, $n = 35$) behaved violently ($M = 0.56$, $SD = 1.13$, range = 0 to 8).

Differences Between Violent and Nonviolent Offenders

Stratified analyses with bivariate logistic regression showed no significant difference between the violent and nonviolent inmates with respect to marital status ($p \leq .63$), criminal record ($p \leq .80$), age at the beginning of incarceration ($p \leq .96$), time spent in the institution ($p \leq .39$), and vocational education ($p \leq .97$).

VRAG Categories

The mean score in the VRAG was 3.65 points, with scores between -23 and $+23$ ($SD = 9.88$). Of offenders, 17% ($n = 18$) were found in VRAG Categories 1 through 3, 15% ($n = 16$) in Category 4, 26% ($n = 28$) in Category 5, 27% ($n = 29$) in Category 6, and 14% ($n = 15$) in Categories 7 through 9 (Table 1).

Predictive Validity of the VRAG

VRAG total score and VRAG categories. A significant effect size could be found for the VRAG total scores and the VRAG categories and verbally aggressive infractions.

Table 1
Number of Offenders per Violence Risk
Appraisal Guide (VRAG) Category

VRAG Categories	<i>n</i>	%	Cumulative (%)
1 through 3	18	16.98	16.98
4	16	15.09	32.08
5	28	26.42	58.49
6	29	27.36	85.85
7 through 9	15	14.15	100
Total	106	100	

Neither the total scores nor the categories were significant predictors for physical violence.

VRAG and different types of offenders. By stratifying the sample between violent offenders ($n = 75$) and sexual offenders ($n = 31$), a significant effect size was found for the VRAG total score of sexual offenders and verbally aggressive infractions; this effect was the only one of the model that retained robust predictive power also under consideration of the alpha parameter. The alpha overdispersion parameter accounts for unexplained variation associated with unobserved predictors. Note that the alpha value was higher than the incidence-rate ratios in all instances, indicating that other unobserved factors are better suited to predict intramural infractions than the VRAG.

No significant effect was found for VRAG total scores of sexual offenders and physically violent infractions. VRAG categories of sexual offenders were not predictive of physically or verbally violent infractions. Likewise, the results for both VRAG total scores and VRAG categories of violent offenders showed no statistical significance for either physically or verbally violent infractions (Table 2).

Discussion

Previous research has shown that the ability of the VRAG to predict extramural violent recidivism is valid across various geographical, cultural, and language areas (Pham, Ducro, Marghem, & Réveillère, 2005; Urbaniok et al., 2006). Several studies have been conducted to assess the predictive validity of the VRAG for in-prison violence (Doyle et al., 2002; Kroner & Mills, 2001). So far no VRAG replication study on intramural violence in the German-speaking part of Europe has been published. The present study aimed to undertake the first evaluation of the accuracy of the VRAG in predicting violent or aggressive behavior of prison inmates in Switzerland.

Table 2
Summary of Negative Binomial Models for Violence
Risk Appraisal Guide (VRAG) Predictions of
Different Types of Intramural Infractions

	VRAG	IRR	SE	95% CI (IRR)	alpha	95% CI (alpha)
VI	Score	1.008	0.022	0.967-1.051	2.235	1.099-4.546
VA	All offenders	1.064*	0.026	1.014-1.116	3.331	1.769-6.271
VI	5 categories	1.125	0.135	0.889-1.423	2.196	1.080-4.465
VA	All offenders	1.391*	0.201	1.049-1.845	3.454	1.856-6.430
VI	Score	1.040	0.056	0.935-1.157	2.132	0.633-7.173
VA	Sexual offenders	1.106*	0.048	1.017-1.204	1.913	0.678-5.400
VI	5 categories	1.389	0.401	0.788-2.447	2.023	0.601-6.810
VA	Sexual offenders	1.494	0.382	0.905-2.466	2.490	0.977-6.348
VI	Score	0.999	0.023	0.954-1.046	2.172	0.882-5.351
VA	Violent offenders	1.012	0.030	0.954-1.072	2.552	0.856-7.609
VI	5 categories	1.043	0.141	0.800-1.359	2.161	0.879-5.317
VA	Violent offenders	1.130	0.187	0.816-1.563	2.517	0.845-7.502

Note: VI = violent infractions: physically aggressive major disciplinary infraction; VA = verbally aggressive, threats; alpha = correction parameter for overdispersion; CI = confidence interval; IRR = incidence-rate ratios.

* $p < .05$.

Our results show, in accordance with the findings of Edens et al. (2006), only a moderate effect between VRAG scores and institutional misconduct. In our sample, hardly any support can be found for extending the use of the VRAG to the prediction of institutional misconduct.

Although all of the information for rating the VRAG scores was solely obtained from file data, the authors are confident that the data provided in the records met the criteria for reliable ratings of the instrument in question. Previous studies have successfully used case file information exclusively, given detailed records, to score VRAG and PCL-R (Grann, Langstrom, Tengstrom, & Kullgren, 1999; Harris et al., 1993).

A significant and robust effect size was found only for VRAG total scores of sexual offenders and verbally violent infractions. There is no obvious theoretical background to explain why the VRAG total scores should predict acts of verbal aggression better for sexual offenders than for violent offenders. This phenomenon will be the subject of further studies.

The study raises questions about the utility of the VRAG in predicting intramural infractions in the particular population investigated. However, the VRAG may be of some use in predicting verbally aggressive behavior of incarcerated sex offenders in the population investigated and thus contribute to some degree to the safety of both prison staff and inmates alike.

References

- Andrews, D. A., & Bonta, J. (2003). *The psychology of criminal conduct* (3rd ed.). Cincinnati, OH: Anderson.
- Bonta, J., Law, M., & Hanson, K. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. *Psychological Bulletin*, *123*, 123-142.
- Doyle, M., Dolan, M., & McGovern, J. (2002). The validity of North American risk assessment tools in predicting in-patient violent behavior in England. *Legal and Criminological Psychology*, *7*, 141-154.
- 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.
- Grann, M. (1998). *Personality disorder and violent criminality: A follow-up study with special reference to psychopathy and risk assessment*. Stockholm: Karolinska Institute, Department of Clinical Neuroscience and Family Medicine.
- Grann, M., Langstrom, N., Tengstrom, 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.
- Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: a meta-analysis. *Psychological Assessment*, *12*, 19-30.
- Harris, G. T., Rice, M. E., & Cormier, C. A. (2002). Prospective replication of the Violence Risk Appraisal Guide in predicting violent recidivism among forensic patients. *Law and Human Behavior*, *26*, 377-394.
- Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behavior*, *20*, 315-335.
- Hilton, N. Z., & Simmons, J. L. (2001). The influence of actuarial risk assessment in clinical judgments and tribunal decisions about mentally disordered offenders in maximum security. *Law and Human Behavior*, *25*, 393-408.
- Horstkotte, H. (1986). Strafrechtliche Fragen zur Entlassungspraxis nach §67d Abs. 2 StGB [Penological questions on the discharge practice according to §67d, Section 2 of the (German) penal code]. *Monatsschrift für Kriminologie und Strafrechtsreform*, *69*, 332-340.
- Jockusch, U. (1990). Therapeutischer Massregelvollzug—Leistungsfähigkeit und Grenzen [Hospital order treatment—Potentials and limitations]. *Recht und Psychiatrie*, *1*, 14-18.
- Krippendorff, K. (1980). *Content analysis. An introduction to its methodology*. Beverly Hills, CA: Sage.
- 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.
- Kühl, J., & Schumann, K. F. (1989). Prognosen im Strafrecht—Probleme der Methodologie und Legitimation [Predictions (of recidivism)—Methodological and legitimate problems]. *Recht und Psychiatrie*, *7*, 126-148.
- Loza, W., & Dhaliwal, G. K. (1997). Psychometric evaluation of the Risk Appraisal Guide (RAG): A tool for assessing violent recidivism. *Journal of Interpersonal Violence*, *12*, 779-793.
- McBride, M. (1999, February). *Predicting violence among federal inmates*. Paper presented at the Corrections Research Forum, Toronto, Canada.
- Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology*, *62*, 783-792.
- Nichols, T. L., Vincent, G. M., Whittemore, K. E., & Ogloff, J. R. P. (1999, November). *Assessing risk of inpatient violence in a sample of forensic psychiatric patients: Comparing the PCL:SV, HCR-20 and VRAG*. Paper presented at Risk Assessment and Risk Management: Implications for the Prevention of Violence, Vancouver, Canada.
- Pham, T. H., Ducro, C., Marghem, B., & Réveillère, C. (2005). Evaluation du risque de récidivie au sein d'une population de délinquants incarcéré ou interné en Belgique francophone [Prediction of recidivism among prison inmates and forensic patients in Belgium]. *Annals Médico Psychologiques*, *163*, 842-845.

- Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.
- Rasch, W. (1994). Die Prognose im Massregelvollzug als kalkuliertes Risiko [The prediction of recidivism in offenders sentenced to hospital order treatment as a calculated risk]. In W. Frisch (Ed.), *Prognoseentscheidungen in der strafrechtlichen Praxis* (pp. 235-251). Baden-Baden, Germany: NOMOS.
- Rice, M. E., & Harris, G. T. (1995). Violent recidivism: Assessing predictive validity. *Journal of Consulting and Clinical Psychology, 63*, 737-748.
- Rice, M. E., & Harris, G. T. (1997). Cross validation and extension of the Violence Risk Appraisal Guide for child molesters and rapists. *Law and Human Behavior, 21*, 231-241.
- Rice, M. E., Harris, G. T., & Cormier, C. A. (1992). An evaluation of a maximum security therapeutic community for psychopaths and other mentally disordered offenders. *Law and Human Behavior, 16*, 399-412.
- Rice, M. E., Harris, G. T., & Quinsey, V. L. (1990). A follow-up of rapists assessed in a maximum-security psychiatric facility. *Journal of Interpersonal Violence, 5*, 435-448.
- Swets, J., Dawas, R., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest, 1*, 1-26.
- Urbaniok, F., Noll, T., Grunewald, S., Steinbach, J., & Endrass, J. (2006). Prediction of violent and sexual offences: A replication study of the VRAG in Switzerland. *Journal of Forensic Psychiatry and Psychology, 17*, 23-31.
- Walters, G. D. (2006). Risk-appraisal versus self-report in the prediction of criminal justice outcomes. *Criminal Justice and Behavior, 33*, 279-304.