

Predicting Reoffence in Sexual Offender Subtypes: A Prospective Validation Study of the German Version of the Sexual Offender Risk Appraisal Guide (SORAG)

Martin Rettenberger, Reinhard Eher

Federal Documentation Centre for Sex Offenders, Austrian Prison System, Gerichtsgasse 6, A-1210 Wien

[Sexual Offender Treatment, Volume 2 (2007), Issue 2]

Abstract

This study is part of a prospective, longitudinal research project to evaluate the reliability and validity of different recidivism risk assessment methods for sexual offenders under community supervision for scientific and practical use in the German-speaking part of Europe. In this paper we present the German adaptation of the Sexual Offender Risk Appraisal Guide (SORAG), a specific risk assessment tool for sexual offenders that was developed and published in 1998 in Canada. We examined interrater reliability, concurrent validity and predictive validity of the German version of the instrument with a sample of 254 male sexual offenders arrested in the Austrian penal system. The German SORAG showed good results for interrater reliability and concurrent validity. The predictive validity was determined using subgroups of the sample based on each offender's index offence. The results were predominantly good, but the ability of the instrument to predict sexual and violent recidivism varied depending on offender type. The best results could be obtained for the prediction of re-imprisonment. The conclusion of this study is that the German version of the SORAG performs equally well and is most useful in predicting serious reoffenses.

Key words: SORAG, risk assessment, sexual offender, validity, recidivism

Introduction

In the last years there has been growing interest in validated risk assessment tools for sex offenders in the German speaking part of Europe (Noll, Endrass, Rossegger, & Urbaniok, 2006). Compared to North America where actuarial risk assessment instruments are regarded as state of the art there exist only a few studies in Germany, Austria or Switzerland that have dealt with these instruments. Although the results of a number of studies from North America support the utility of the actuarial approach for the risk assessment for sexual offenders (e.g. Barbaree, Seto, Langton, & Peacock, 2001), validation studies of these instruments also on German-speaking populations are required before giving a general recommendation for the use of these instruments in Germany and Austria.

In this study we present the results of a cross-validation of the German adaptation of the *Sexual Offender Risk Appraisal Guide* (SORAG; Quinsey, Harris, Rice, & Cormier, 2006; Rettenberger & Eher, 2007) and its utility to predict relapse in a prospective longitudinal design of sexual offenders

under community supervision. The SORAG is an actuarial risk assessment tool for sexual offenders and was developed by the Canadian forensic researcher Vernon L. Quinsey and his colleagues. This instrument is a modification of the *Violence Risk Appraisal Guide* (VRAG) that was developed to predict violent and sexual recidivism among male offenders; 10 of the 14 items of the SORAG are the same items as in the VRAG. The SORAG is conceptualized for the risk assessment for sexual offenders to assess the likelihood of violent recidivism which includes sexual offences involving physical contact with the victim.

The instrument consists of 14 weighted items (see Table 1). After scoring these the evaluator adds up the item scores and gets the total score of the SORAG. Based on the total score the evaluator can allocate the offender to one of nine risk categories. By means of these risk categories it is possible to infer to empirically calculated probabilities of violent (including sexual) recidivism after seven and ten years, respectively.

Table1: Items and basic coding rules of the SORAG (Quinsey et al., 2006)

Item Number	Risk Factor	Coding Rule
1	Lived with both biological parents to age 16 (except for death of parent) – Score <i>no</i> if offender did not live continuously with both biological parents until age 16, except if one or both parents died. In case of parent death, score as for yes	Yes = -2 No = +3
2	Elementary school maladjustment (up to and including Grade 8)	No problems = -1 Slight or moderate problems = +2 Severe problems = +5
3	History of alcohol problems – Allot one point for each of the following: alcohol abuse in biological parent, teenage alcohol problem, adult alcohol problem, alcohol involved in a prior offense, alcohol involved in the index offense	0 = -1 1 or 2 = 0 3 = +1 4 or 5 = +2
4	Marital status (or lived common law in the same home for at least 6 months) – At time of index offense	Ever married = -2 Never married = +1
5	Criminal history score for convictions and charges for nonviolent offenses prior to the index offense (from the Cormier-Lang system)	Score 0 = -2 Score 1 or 2 = 0 Score of 3 or above = +3
6	Criminal history score for convictions and charges for violent offenses prior to the index offense (from the Cormier-Lang system)	Score 0 = -1 Score 2 = 0 Score of 3 or above = +6

7	Number of convictions for previous sexual offenses (pertains to convictions for sexual offenses that occurred prior to the index offense) – Count any offenses known to be sexual, including, for example, indecent exposure	0 = -1 1 or 2 = +1 ≥ 3 = +5
8	History of sex offenses against girls under age 14 <i>only</i> (includes index offense; if offender was less than 5 years older than victim, always score +4)	Yes = 0 No = +4
9	Failure on prior conditional releases (includes parole violation or revocation; breach of or failure to comply with recognizance or probation; bail violation; and any new charges, including the index offense, while on a conditional release)	No = 0 Yes = +3
10	Age at index offense (at most recent birthday)	≥ 39 = -5 34-38 = -2 28-33 = -1 27 = 0 ≤ 26 = +2
11	Meets DSM-III criteria for any personality disorder	No = -2 Yes = +3
12	Meets DSM-III criteria for schizophrenia	Yes = -3 No = +1
13	Phallometric test results	<i>All</i> indicate nondeviant sexual preferences = -1 <i>Any</i> test indicates deviant sexual preferences = +1
14	Hare Psychopathy Checklist – Revised score (PCL-R; Hare, 1991)	≤ 4 = -5 5-9 = -3 10-14 = -1 15-24 = 0 25-34 = +4 ≥ 35 = +12

In order to use the SORAG in Austria, we had to adapt the original coding guidelines to the forensic context in this country (Rettenberger & Eher, 2007). This concerns especially the items number 11 and 12, because of changes of the diagnostic criteria in the new version of the DSM (American Psychiatric Association [2003], APA). A further important modification concerns the item number 13: In the German-speaking part of Europe phallometric evaluations are not conducted routinely, so this information is simply not available¹. Therefore, a diagnosis of sexual deviance according to DSM or ICD criteria replaced the phallometric evidence. If a sexual offender met the DSM or ICD criteria for any paraphilia the evaluator had to score +1, otherwise he had to score -1.

There are no published validation studies of the SORAG in the German-speaking part of Europe until now². Most of the validation studies derive from Anglo-American countries. Harris, Rice, Quinsey, Lalumère, Boer, and Lang (2003) reviewed the results of these studies and concluded that several studies (e.g., Barbaree, Seto, Langton, & Peacock, 2001; Nunes, Firestone, Bradford,

Greenberg, & Broom, 2002; Rice & Harris, 2002) have shown high accuracy of the SORAG in the prediction of violent (including sexual) recidivism and moderate accuracy in predicting sexual offenses. For the prediction of violent (including sexual) recidivism they calculated a median AUC value of AUC = .75.

Although the results of the predictive accuracy of the SORAG has been reasonably consistent across studies, Bartosh, Garby, Lewis, and Gray (2003) suggested that the predictive validity of the instrument varied depending on the type of sexual offender. According to these authors, the SORAG could significantly predict sexual, violent, and overall recidivism for extrafamilial child molesters (AUC values ranged from .70 to .93) and for incest offenders (AUC ranged from .72 to .91). As regards to rapists and hands-off offenders, however, the SORAG showed much lower predictive power (AUC ranged from .46 to .71). Ducro and Pham (2006), retrospectively, evaluated the predictive accuracy of the SORAG on Belgian sexual offenders committed to a forensic facility. For the total sample the instrument showed strong predictive validity for general (AUC = .70) and violent (AUC = .72) recidivism and moderate predictive validity for sexual recidivism (AUC = .64). Depending on offender subgroup and recidivism criterion the AUC values ranged from .64 to .77. The results of Bartosh et al. (2003) and Ducro & Pham (2006) support the evidence that the SORAG shows good predictive validity, whereas the results varied depending on sex offender type.

The primary aim of this prospective, longitudinal study was to investigate the usefulness and the predictive validity of the German version of the SORAG in the forensic context of a German speaking country. Therefore, we present data on the interrater reliability, the concurrent validity, and the predictive validity of this German version. Furthermore, the predictive validity was also investigated in subgroups of the sample based on offender's index offence. On the one hand, we expected the SORAG to show high predictive accuracy and that our findings would be in accordance with previous studies from North America and Europe. On the other hand, we expected differences in the predictive accuracy of the SORAG depending on sexual offender subgroup and recidivism criterion.

Method

Subjects

Two hundred and fifty-four male sex offenders were investigated in 2002 and 2003 at the *Federal Documentation Centre of Sex Offender of the Austrian Prison System* and followed up after prison release until December, 31, 2005 (mean age at time of release = 40.92 years, SD = 12.53, range 18-72). The follow-up time periods ranged from 24.27 to 66.67 months (M = 38.82 months; SD = 9.51; 24 months was defined to be a minimal follow-up period). The group consisted of 132 child molesters (58 extrafamilial molest offenders and 74 incest offenders), and 120 rapists; one offender was convicted because of sexual burglary and one because of child pornography. The mean duration of imprisonment for the index offense was M = 32.37 months (SD = 22.14), 20.9% (n = 53) had at least committed one previous sexual offense (M = 0.54; SD = 1.80, range 0-16), and 58.7% (n = 149) had at least one prior conviction for any offense (M = 3.88, SD = 6.14, range 0-37).

Of the 254 offenders, 10.2% (n = 26) failed compulsory school or attended school for mentally or physically handicapped children, 34.3% (n = 87) completed school, but had no professional education, 49.6% (n = 126) completed technical school or professional education, 3.1% (n = 8) had a university-entrance diploma or completed technical college, and 0.8% (n = 2) had a university degree. Regarding the level of education, there were no information on 2.0% (n = 5) of offenders.

Procedure and Statistical Analysis

The interrater reliability of the SORAG was examined by the Intraclass–Correlation Coefficient (ICC). The predictive ability of the instrument was analyzed using Pearson correlations and the area under the curve (AUC) of the ROC. Referring to Cohen (1992), Dahle, Schneider, and Ziethen (2007) formulated the following criteria for the classification of the predictive accuracy of risk assessment tools: AUC values of .72 or above ($r \geq .37$) are classified as good and AUC values between .64 and .71 ($r \geq .24$) are classified as moderate. Significant AUC values that are below the value of .64 ($r < .24$) are classified as small³.

In order to establish the concurrent validity of the instrument, we calculated Pearson correlations between the SORAG on the one hand and the *Sexual Violence Risk-20* (SVR-20; Boer, Hart, Kropp, & Webster, 1997), the *Static-99* (Hanson & Thornton, 1999), and the *Psychopathy Checklist-Revised* (PCL-R; Hare, 1991) on the other hand⁴. We chose the SVR-20, SORAG, and PCL-R, because there exists already a number of validation studies (e.g. Rettenberger & Eher, 2006a; Barbaree et al., 2001; Harris, Rice, et al., 2003), so these instruments represent internationally recognized risk assessment tools for sex offenders (e.g. Müller-Isberner, Cabeza, & Eucker, 2000; Habermeyer & Herpertz, 2005). For Pearson correlations and ICC values we use the raw scores of each instrument.

The SVR-20 is a structured clinical guideline designed for the assessment of risk for sexual violence in adult sex offenders. The instrument was developed from a thorough consideration of the empirical literature and the clinical expertise of a number of clinicians. The SVR-20 consists of 20 items, divided into three domains: Psychosocial adjustment, Sexual offenses, and Future plans, which have to be coded by an experienced forensic clinician. Although originally designed as a structured clinical guideline, it is not uncommon for research purposes to add up the items. In this case, the instrument becomes a conceptual actuarial measure (Hanson & Morton-Bourgon, 2007).

The Static-99 is a brief actuarial instrument for the assessment of risk for sexual and violent recidivism of adult sexual offenders. This instrument is derived from the fusion of two previously developed risk assessment instruments, the *Rapid Risk Assessment of Sexual Offender Recidivism* (RRASOR; Hanson, 1997) and a shorter version of the *Structured Anchored Clinical Judgement* (SACJ-Min; Grubin, 1998) composed of ten mainly static risk factors. In terms of the Static-99, there exists already a German adaptation (Rettenberger & Eher, 2006b) of the revised version of this instrument (Harris, Phenix, Hanson, & Thornton, 2003). The German version of the Static-99 showed high predictive accuracy and good results for interrater reliability and concurrent validity (Rettenberger & Eher, 2006a).

In contemporary research and clinical practice, Hare's Psychopathy Checklist-Revised (PCL-R) is the psycho-diagnostic tool most commonly used to assess psychopathy. The PCL-R is based on semistructured interviews and review of file information. Participants are assigned to ratings of 0 (absent), 1 (some indication), or 2 (present) on each of the 20 PCL-R items, tapping characteristics such as impulsivity, irresponsibility, and callousness. Scale scores are obtained by summing the ratings, for a total possible score of 40. The conventional cutoff for making a diagnosis of psychopathy in North America is 30, whereas in Europe the cutoff is mostly 25 (Hart & Hare, 1997; Hartmann, Hollweg, & Nedopil, 2001). In the development of the PCL-R, Hare was specifically interested in the construction of an instrument to quantify psychopathic personality traits, not an instrument to predict reoffenses. Nevertheless, research shows that the PCL-R does a reasonable job in predicting recidivism (e.g. Quinsey, Rice, & Harris, 1995).

In order to establish the interrater reliability, three randomly selected colleagues, that have to be

theoretically and practically skilled in the application of actuarial risk assessment methods, rated independently ten cases. These ten cases were randomly selected within the database of the *Federal Documentation Centre of Sex Offender of the Austrian Prison System*. The second author gave the raters official file information of each case to code the actuarial instruments (e.g. charges, convictions, official criminal records, etc.).

Data on recidivism were retrieved from the official Federal Central Register of the Austrian Ministry of Internal Affairs, whereas each new conviction (with or without imprisonment), that was listed on the official criminal record, was counted as a reoffense. Furthermore, we counted separately each new imprisonment. We defined the following recidivism criteria: Sexual recidivism (with or without imprisonment), nonsexual violent recidivism (with or without imprisonment), general recidivism (with or without imprisonment), and violent recidivims (including sexual and non sexual violent offenses, with and without imprisonment)⁵.

Results

Reconviction Rates

Offenders (N = 254) were categorized as rapists (47.2%; n = 120), extrafamilial child molesters (22.8%; n = 58), and incest offenders (29.1%; n = 74). Two offenders (0.8%) could not be allocated to either of these groups and therefore were excluded. The reconviction rates for the total sample and the three subgroups are shown in Table 2.

Table 2: Reconviction rates of the total sample and each category of index offense type

Recidivism Criterion	Total Sample (N = 254)	Rapists (n = 120)	Extrafamilial Molest Offenders (n = 58)	Incest Offenders (n = 74)
Any Recidivism (Conviction)	25.2% (n = 64)	33.3% (n = 40)	20.7% (n = 12)	16.2% (n = 12)
Any Recidivism (Imprisonment)	15.0% (n = 38)	22.5% (n = 27)	15.5% (n = 9)	2.7% (n = 2)
Sexual Recidivism (Conviction)	3.5% (n = 9)	1.7% (n = 2)	12.1% (n = 7)	0% (n = 0)
Sexual Recidivism (Imprisonment)	3.1% (n = 8)	1.7% (n = 2)	10.3% (n = 6)	0% (n = 0)
Violent non-sexual Recidivism (Conviction)	12.2% (n = 31)	20.8% (n = 25)	1.7% (n = 1)	6.8% (n = 5)
Violent non-sexual Recidivism (Imprisonment)	6.7% (n = 17)	12.5% (n = 15)	1.7% (n = 1)	1.4% (n = 1)

Sexual and/or Violent Recidivism (Conviction)	15.4% (n = 39)	21.7% (n = 26)	13.8% (n = 8)	6.8% (n = 5)
Sexual and/or Violent Recidivism (Imprisonment)	9.4% (n = 24)	13.3% (n = 16)	12.1% (n = 7)	1.4% (n = 1)

Interrater Reliability

Using the critical values for ICC (single measure) of Fleiss (1986; $ICC \geq .75$ = excellent; $.60 \leq ICC < .75$ = good; $.40 \leq ICC < .60$ = moderate; $ICC < .40$ = poor), the interrater reliability of the German version of the SORAG was excellent: We calculated an $ICC = .93$ (single measure, $p < .001$). Furthermore, we also evaluated the interrater reliability of the instruments used for the concurrent validity analyses. All instruments showed excellent interrater reliability: for the Static-99 we calculated an $ICC = .98$, for the SVR-20 an $ICC = .84$, and for the PCL-R = $.93$ (all $p < .001$).

Concurrent Validity

The SORAG showed good concurrent validity: The instrument correlated significantly ($p < .001$) with the Static-99, SVR-20, and PCL-R: The correlation coefficients were $r = .71$, $r = .78$ and $r = .79$, respectively.

Predictive Validity

The mean total score of the SORAG for the 254 sex offenders was 5.59 (SD = 13.49). Regarding the SORAG risk categories, 12.6% (n = 32) were categorized into the first risk group (SORAG total score of ≤ -10), 17.3% (n = 44) into the second risk group (-9 to -4), 17.0% (n = 43) into the third risk group (-3 to +2), 12.2% (n = 31) into the fourth risk group (+3 to +8), 11.4% (n = 29) into the fifth risk group (+9 to +14), 10.6% (n = 27) into the sixth risk group (+15 to +19), 9.8% (n = 25) into the seventh risk group (+20 to +24), 6.3% (n = 16) into the eighth risk group (+25 to +30), and 2.8% (n = 7) into the ninth risk group (≥ 31).

The values of the predictive validity (AUC values of the ROC and Pearson correlations) of the SORAG for the total sample and the three subgroups are shown in Table 3. For the total sample the AUC values are at least .73 and for all recidivism criteria of the total group the AUC and correlation values reach statistical significance. For the rapist subgroup the SORAG shows also significant predictive accuracy, comparable to the values of the total sample. However, the results for sexual recidivism (where just two relapses occurred) for the rapist subgroup reach no statistical significance because of too few cases of relapse within this category. With exception of the results for violent nonsexual recidivism (where just one relapse occurred) the results for the extrafamilial child molester group are at least $AUC = .70$, all $p < .001$. For the incest offender subgroup the recidivism base rates are altogether very low, so the results reach statistical significance only for the any recidivism -criteria. Overall, the AUC values for the imprisonment -criteria are consistently higher than for the conviction -criteria.

Table 3: Predictive validity of the SORAG for the total sample and each category of index offense type

Recidivism Criterion	Total Sample (N = 254)	Rapists (n = 120)	Extrafamilial Molest Offenders (n = 58)	Incest Offenders (n = 74)
Any Recidivism (Conviction)	AUC = .73** (r = .35**)	AUC = .73** (r = .37**)	AUC = .70* (r = .26)	AUC = .62 (r = .26*)
Any Recidivism (Imprisonment)	AUC = .83** (r = .41**)	AUC = .80** (r = .41**)	AUC = .76* (r = .30*)	AUC = .95* (r = .31**)
Sexual Recidivism (Conviction)	AUC = .73* (r = .14*)	AUC = .71 (r = .09)	AUC = .76* (r = .28*)	-- ^a
Sexual Recidivism (Imprisonment)	AUC = .73* (r = .14*)	AUC = .71 (r = .09)	AUC = .76* (r = .26)	-- ^a
Violent non-sexual Recidivism (Conviction)	AUC = .75** (r = .29**)	AUC = .72** (r = .31**)	AUC = .56 (r = .01)	AUC = .63 (r = .21)
Violent non-sexual Recidivism (Imprisonment)	AUC = .83** (r = .30**)	AUC = .80** (r = .34**)	AUC = .56 (r = .01)	AUC = .92 (r = .18)
Sexual and/or Violent Recidivism (Conviction)	AUC = .76** (r = .33**)	AUC = .73** (r = .33**)	AUC = .74* (r = .27*)	AUC = .63 (r = .21)
Sexual and/or Violent Recidivism (Imprisonment)	AUC = .82** (r = .33**)	AUC = .81** (r = .36**)	AUC = .74* (r = .24)	AUC = .92 (r = .18)

* p < .05, ** p < .001, ^a no reoffenses

Discussion

The present study is the first assessing the validity of a German version of the SORAG in a prospective design. The results show a normal distribution of the total scores and exhibited the mean score of the SORAG to be comparable to results of previous studies (e.g. Bartosh et al., 2003; Barbaree et al., 2001). This indicates that our sample is representative for a general sexual offender population. Regarding the critical values of ICC (single measure) of Fleiss (1986), interrater reliability of the German version of the instrument is excellent. These results are also comparable to previous studies (e.g. Ducro & Pham, 2006; Bartosh et al., 2003). Furthermore, the values of the concurrent validity analyses can be interpreted as satisfactory (Harris, Rice, et al., 2003). Although the PCL-R is not designed as an actuarial measure for sex offender risk assessment, the high correlation between SORAG and PCL-R is not surprising: The PCL-R is included in the SORAG, so the PCL-R total score has a direct influence on the SORAG total score. The high correlations between the SORAG on the one hand and Static-99 and SVR-20 on the other hand reveal the fact that that all 3 instruments are designed to predict sexual recidivism partially using the same items (e.g. criminal history). But there are, however, less than perfect associations between these instruments, so it has to be one aim of further research on this topic to clarify similarities and differences between actuarial methods⁷.

Referring to the critical value descriptors for AUC values by Dahle, Schneider, and Ziethen (2007), the predictive validity of the German SORAG for the total sample can be classified as good. The AUC values for the imprisonment-criteria were consistently higher than for the conviction-criteria. This result is an interesting finding with importance for forensic practitioners, since it suggests the instrument is better at predicting more serious relapses (which are punished by a prison sentence).

Originally, the SORAG was developed to predict nonsexual violent and/or sexual recidivism (Quinsey et al., 2006). Therefore, the recidivism criterion is of particular interest and the present study showed good results for this criterion (AUC = .76 for new convictions and AUC = .82 for new imprisonments). Even more importantly, these AUC values are slightly higher than those from previous results of Anglo-American studies (e.g. Barbaree et al., 2001; Harris, Rice, et al., 2003)⁶.

The German version of the SORAG tends to yield more accurate results for rapists than for both pedosexual subgroups. With the exception of the results for sexual recidivism, the instrument shows good predictive accuracy for the rapist subgroup. Although the AUC value for sexual recidivism can be classified as moderate, there are too few cases of relapse within this category in order to interpret these results reasonably (which is also reflected by the fact that this ROC is not significant). For the extrafamilial child molester group the results are similar: With exception of the results for violent nonsexual recidivism where just one relapse occurred the predictive validity of the SORAG can be interpreted as satisfactory.

As mentioned before, it is important to take a more precise look at the recidivism base rates and the follow-up time periods when interpreting the results of the sexual offender subgroup analyses: During the follow-up time period in our study there are subgroups and recidivism criteria with no (incest offender subgroup with no sexually motivated relapses) or very few relapses (only one nonsexually motivated violent relapse in the extrafamilial child molester group). In these recidivism categories the SORAG shows the lowest predictive accuracy. In the innerfamilial child sexual abuse group no sexual relapses occurred. Therefore, no AUC values could be calculated. When taking into account the low recidivism rates in some categories of some offender subgroups, one has to address one major limitation of this study: the small sample sizes, even though AUC measures usually are seen to be less sensitive to low base rates than other statistical procedures. Given to our results one can expect that predictive accuracy of the instrument will be demonstrated even for these subgroups and recidivism criteria after a longer at-risk period.

To conclude, our results support the utility of actuarial risk assessment methods like the SORAG even in non English speaking countries. The results of the German version to predict relapse in sexual offenders are predominantly good. Nevertheless, the ability to predict recidivism varied depending on offender type and relapse category with a consistently better performance on more serious reoffenses. The results of the present study again indicate the helpfulness of actuarial instruments like the SORAG for assessing the recidivism risk in sexual offenders.

References

1. American Psychiatric Association (2003). Diagnostisches und statistisches Manual psychischer Störungen: Textrevision DSM-IV-TR. Göttingen: Hogrefe.
2. Barbaree, H. E., Seto, M. C., Langton, C. M., & Peacock, E. J. (2001). Evaluating the Predictive Accuracy of Six Risk Assessment Instruments for Adult Sex Offender. *Criminal*

- Justice and Behavior, 28, 490-521.
3. Bartosh, D. L., Garby, T., Lewis, D., & Gray, S. (2003). Differences in the Predictive Validity of Actuarial Risk Assessments in Relation to Sex Offender Type. *International Journal of Offender Therapy and Comparative Criminology*, 47, 422-438.
 4. Boer, D. P., Hart, S. D., Kropp, P. R., & Webster, C. D. (1997). *Manual for the Sexual Violence Risk - 20: Professional Guidelines for Assessing Risk of Sexual Violence*. Vancouver/CA: The Mental Health, Law, & Policy Institute.
 5. Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, 112, 155-159.
 6. Dahle, K. P., Schneider, V., & Ziethen, F. (2007). Standardisierte Instrumente zur Kriminalprognose. *Forensische Psychiatrie, Psychologie, Kriminologie*, 1, 15-26.
 7. Douglas, K. S. (2001). HCR-20 Violence Risk Assessment Scheme: Overview and Annotated Bibliography. <http://www.sfu.ca/psychology/groups/faculty/hart/violink.htm>.
 8. Ducro, C. & Pham, T. (2006). Evaluation of the SORAG and the Static-99 on Belgian Sex Offenders Committed to a Forensic Facility. *Sexual Abuse: A Journal of Research and Treatment*, 18, 15-26.
 9. Fleiss, J. L. (1986). *The Design and Analysis of Clinical Experiments*. New York: Wiley.
 10. Grubin, D. (1998). *Sex Offending Against Children: Understanding the Risk*. Police Research Series, Paper 99. London: Home Office.
 11. Habermeyer, E. & Herpertz, S. (2005). Bedeutung des "psychopathy"-Konzeptes von Hare für Sexualstraftaten. In D.Schläfke, F. Häbeler, & J. M. Fegert (Eds.), *Sexualstraftaten* (pp. 131-141). Stuttgart: Schattauer.
 12. Hanson, R. K. (1997). *The Development of a Brief Actuarial Risk Scale for Sexual Offense Recidivism*. User Report 1997-04. Ottawa: Department of the Solicitor General of Canada.
 13. Hanson, R. K. & Thornton, D. (1999). *Static-99: Improving Actuarial Risk Assessments for Sex Offenders*. User Report 99-02. Ottawa: Department of the Solicitor General of Canada.
 14. Hanson, R. K. & Morton-Bourgon, K. E. (2007). *The Accuracy of Recidivism Risk Assessment for Sexual Offenders: A Meta-Analysis*. User Report 07-01. Ottawa: Department of the Solicitor General of Canada.
 15. Hare, R. D. (1991). *Manual for the Psychopathy-Checklist-Revised*. Toronto: Multi Health Systems.
 16. Harris, G. T., Rice, M. E., Quinsey, V. L., Lalumière, M. L., Boer, D., & Lang, C. (2003). A Multisite Comparison of Actuarial Risk Instruments for Sex Offenders. *Psychological Assessment*, 15, 413-425.
 17. Harris, G. T., Phenix, A., Hanson, R. K., & Thornton, D. (2003). *Static-99 Coding Rules Revised - 2003*. Ottawa: Department of the Solicitor General of Canada.
 18. Hart, S. D. & Hare, R. D. (1997). *Psychopathy: Assessment and Association with Criminal Conduct*. In D.Stoff, J. Breiling, & J. D. Maser (Eds.), *Handbook of Antisocial Behavior* (pp. 22-35). New York: Wiley.
 19. Hartmann, J., Hollweg, M., & Nedopil, N. (2001). Quantitative Erfassung dissozialer und psychopathischer Persönlichkeiten bei der strafrechtlichen Begutachtung. *Nervenarzt*, 72, 365-370.
 20. Marshall, W. L. (2005). Grenzen der Phallometrie. *Recht & Psychiatrie*, 23, 11-23.
 21. Müller-Isberner, R., Cabeza, S. G., & Eucker, S. (2000). Die Vorhersage sexueller Gewalttaten mit dem SVR-20. Haina: Institut für Forensische Psychiatrie.
 22. Noll, T., Endrass, J., Rossegger, A., & Urbaniok, F. (2006). Die Risikokalkulation für die Begehung von Gewalt- und Sexualstraftaten mit Hilfe des Static-99. *Monatsschrift für Kriminologie und Strafrechtsreform*, 89, 24-33.
 23. Nunes, K. L., Firestone, P., Bradford, J. M., Greenberg, D. M., & Broom, I. (2002). A Comparison of Modified Versions of the Static-99 and the Sex Offender Risk Appraisal Guide (SORAG). *Sexual Abuse: A Journal of Research and Treatment*, 14, 253-269.

24. Quenzer, C. (2005). Prädiktive Validität ausgewählter aktuarischer Verfahren bei Sexualdelinquenten. Berlin: Unpublished Diploma Thesis.
25. Quinsey, V. L., Rice, M. E., & Harris, G. T. (1995). Actuarial Prediction of Sexual Recidivism. *Journal of Interpersonal Violence*, 10, 85-105.
26. Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. (1998). *Violent Offenders: Appraising and Managing Risk*. Washington, DC: American Psychological Association.
27. Rettenberger, M. & Eher, R. (2006a). Actuarial Assessment of Sex Offender Recidivism Risk: A Validation of the German Version of the Static-99. *Sexual Offender Treatment*, 1, 1-11.
28. Rettenberger, M. & Eher, R. (2006b). Die revidierten Kodierungsrichtlinien des Static-99. <http://www.vffw.org>.
29. Rettenberger, M. & Eher, R. (2007). Aktuarische Kriminalprognose bei Sexualstraftätern: Die deutsche Version des Sex Offender Risk Appraisal Guide (SORAG). <http://www.vffw.org>.
30. Rice, M. E. & Harris, G. T. (2002). Men who Molest their Sexually Immature Daughters: Is a Special Explanation Required? *Journal of Abnormal Psychology*, 111, 329-339.
31. Vogel, V. d., Ruiter, C. d., Beek, D. v., & Mead, G. (2004). Predictive Validity of the SVR-20 and Static-99 in a Dutch Sample of Treated Sex Offenders. *Law and Human Behavior*, 28, 251.

Footnotes

¹ Furthermore, the evidence of the reliability and criterion validity of phallometric testing in common seems to be weak, although research has suggested a limited value in predicting subsequent recidivism. Therefore Marshall (2005) concluded that the routine use of phallometric assessments as part of the evaluation of sexual offenders can not be recommended.

² However, there are unpublished validation studies in Germany that indicate good predictive validity results (e.g. Quenzer, 2005).

³ De Vogel, de Ruiter, van Beek, and Mead (2004) use with reference to Douglas (2001) the following interpretation of critical AUC values: AUC values of .70 and above are classified as moderate, and values above .75 are classified as good.

⁴ The upcoming task of the risk assessment research project of the Federal Documentation Centre of Sex Offender of the Austrian Prison System is the advanced analysis of other risk assessment tools (e.g. the SVR-20, Static-99, and PCL-R) and the comparison of the predictive accuracy of these instruments.

⁵ The SORAG was developed to predict violent or/and sexual reoffenses (Quinsey et al., 2006).

⁶ In this context, it should be noted that the small modifications of the German coding rules could be one reason for differences in the results of the predictive accuracy. On the other hand these results show that the German adaptation of the SORAG works at least as good as the original version: This indicates, for example, that a clinical diagnoses of paraphilia represents a good substitution for phallometric assessment results.

⁷ For example, it is possible that the instruments measure different dimensions of recidivism risk. In this study we were not able to focus on this issue, but future research may provide more precise findings about this topic.

Author address

*Reinhard Eher, M.D., Associate Professor
Federal Documentation and Evaluation Centre for Violent and Sexual Offenders
Gerichtsgasse 6
A-1210 Vienna
Austria
E-Mail:  reinhard.eher@justiz.gv.at
Phone: ++43 1 27 85 300 31
Fax: ++43 1 27 85 300 32*