



I'm not robot



Continue

## Stable 2007 scoring manual

Sexual offenders cause serious harm to society which causes justified public concern. Over the past 10 years, psychology has developed the ability to reliably classify male sex offenders as low, moderate, and high risk for sexual relapse (Minnesota Sex Offender Screening Tool-Revised, Risk Matrix-2000, Rapid Risk Assessment for Sexual Offense Recidivism, and STATIC-99) based on historical, static, and immutable risk factors. The static structure of these tests effectively precludes their ability to measure risk changes. The STABLE-2007 and ACUTE-2007 tools are specialized tools designed to evaluate and track changes in risk status over time by evaluating modifiable dynamic risk factors. Stable dynamic risk factors are personal skill deficits, predilections, and learned behaviors that are correlated with sexual relapse but can be modified through an effort intervention process (i.e., treatment or supervision). If such intervention is carried out in such a way as to reduce these risk-relevant factors, there would be a concomitant reduction in the likelihood of sexual recurrence. Acute dynamic risk factors are highly transient conditions that last only hours or days. These factors are rapidly changing stress, environmental and intrapersonal conditions or events that have been proven by previous research as related to an impending sexual re-offense. Such tools should be used to inform correctional managers about the amount of risk they manage, inform decisions about community levels of treatment and supervision, and estimate changes in risk status before and after treatment or other interventions. In the late 1990s, Karl Hanson and Andrew Harris began investigating the relationship between sexual relapse and dynamic, modifiable, and sexual recidivism-related risk factors. This work produced the SONAR (Sex Offender Needs Assessment Rating), which demonstrated adequate internal consistency and a moderate ability to differentiate sexual repeat offenders from non-repeat offenders. Extending this work, Hanson and Harris divided SONAR into two parts, creating a stable dynamic risk measure, STABLE-2000 (16 items) and an acute measure of dynamic risk, ACUTE-2000 (8 items). To test these new tools, Hanson and Harris launched a prospective study, the Dynamic Supervision Project, which involved every Canadian province and territory and the states of Alaska and Iowa in a robust test of risk assessment methodologies. A total of 156 probation and probation officers completed repeated risk assessments at (static, stable and acute) levels on 997 sex offenders in 16 jurisdictions. All probation and probation officers who marked the risk of reoffending for these community-based sex offenders were trained in assessing the risk of sexual offenders by participating in a 2-day scoring-focused workshop examples of cases. Information of sexual, violent and other recidivism was collected from official criminal stories after a median of 41 months of follow-up. The results showed that both STABLE-2000 and ACUTE-2000 added higher predictive validity than only STATIC-99 demonstrated. The rate of sexual relapse for this widely disparate group of community-based sex offenders was 7.6% after 3 years (n = 790). Empirical changes in scoring have been recommended and this research has led to the development of two best dynamic risk measures, the STABLE-2007 and ACUTE-2007 instruments. STABLE-2007 assesses 13 stable risk factors that have been shown to be related to sexual recidivism: significant social influences, ability to relate, emotional identification with children, hostility to women, general social rejection, lack of concern for others, impulsivity, poor problem-solving skills, negative emotionality, sexual desire and concerns, sex such as coping, deviant sexual preference, and cooperation with supervision. Each of these 13 elements is evaluated on a 3-point scale (0 = no obvious problem, 1 = some obvious problem, and 2 = obvious significant problem) for a total of 26 possible points. Emotional identification with children is not marked for those offenders who do not have a child victim, and the scale is then marked on 24 points for that group. The offender's STATIC-99 score is then combined with his STABLE-2007 score to produce percentage estimates of sexual reoffending, sexual recidivism plus sexual assault, violent recidivism, any criminal recidivism (excluded violations), and any criminal recidivism including violations at age 1, 2, 3, and 4. ACUTE-2007 assesses seven rapidly evolving acute risk factors related to sexual relapse. On this scale, there are two factors. The first factor involves sexual and violent reoffending and uses the following four risk factors: victim access, hostility, sexual concern, and refusal of supervision. The second factor involves general criminal recidivism using the four factors mentioned above plus emotional collapse, collapse of social media, and substance abuse for a total of seven elements. Each of these seven elements is evaluated on a scale of 4 points (0 = no obvious problem, 1 = some obvious problem, 2 = obvious significant problem, and 3 = intervene now) for a total of 14 possible points. An Intervene score now requires immediate action to prevent imminent re-offending or supervisory catastrophes such as suicide. Once the ACUTE-2007 score is obtained, this result is combined with the score STABLE-2007 of the offender to estimate an overall risk priority. The offender is nominally classified as a low, moderate or high risk of sexual and violent reoffending and as a low, moderate or high risk of general criminal reoffending. Appropriate based risk reports can then be applied to determine the priority for action. The STABLE-2007 and ACUTE-2007 instruments are easier to score than their predecessors, and combinations of the STATIC-99 and STABLE-2007 instruments have produced receiver operational characteristic curve (ROC) values for sexual re-offense commonly in the 0.76 range. When used by conscientious officers, the combined STATIC-99/STABLE-2007 scores produced a SEXUAL RECIDIVISM ROC of 0.84 and 0.80 for violent reoffending. The STABLE-2007 and ACUTE-2007 assessments added more predictive power than is only available for static risk assessments. This study provides additional evidence that trained community oversight officers can obtain valid and useful risk assessments for sexual offenders. Results of this kind, even taking into account the need for cross-replication and validation, suggest significant political and practical implications for community supervision of sexual offenders. The STABLE-2007 and ACUTE-2007 tools have proven predictive validity beyond that of the SONAR and STABLE-2000/ACUTE-2000 packages. STABLE-2007 and ACUTE-2007 are available free of charge from the authors. The authors no longer support or recommend the use of SONAR, STABLE-2000, or ACUTE-2000, but recommend STABLE-2007 and ACUTE-2007 to assess dynamic changes in risk for sexual offenders. Reference: Hanson, R. K., Harris, A. J. R., Scott, T.-L., & Helmus, L. (2007). Assessment of the risk of sexual offenders on community supervision: The dynamic supervision project (user report, corrections research). Ottawa, ON, Canada: Public Safety Canada. Recovered See also: The Stable-2007 is a tool that measures potentially editable but bearable factors for months or years. Stable 2007 consists of a guided interview process covering 13 main areas of stable risk in the offender, namely: Significant capacity for social influences for stable relationships Emotional identification with children Hostility towards women General social rejection Lack of concern for others Impulsive Poor Problem Solving Skills Emotionality Negative Sex Drive / Sex Concern Sex as Coping Deviant Sexual Preferences Cooperation with supervisor As with Static-99. When you choose to create a new Stable-2007 assessment, you can specify the type of offense and purpose of the assessment (initial or reevaluation) in the General Information section and specify the other fields in the score form. The scores for each question are determined by the compiler during the The evaluation screen also allows the compiler to insert notes for specific questions, which are saved as part of the evaluation record. The STABLE 2007 score report calculates the total score and risk category and presents them together with the original answers in a concise, easy-to-read format suitable for printers: ISSN 1862-2941 Online-Issue: # 2-2017 = Fernandez A field examination of the inter-rater reliability of the Static-99 and STABLE-2007 marked by the officers of the Yolanda M correctional program. Fernandez1, L. Maake Helmus2, L. Therapeutic Presences of Canada2Saskatchewan University (Treatment of Sexual Offenders, Volume 12 (2017), Number 2) Abstract Objective: For measures to be considered useful they must be: 1) reasonably standardized; 2) valid; 3) reliable. The third of these principles, reliability, is the extent to which the results/scores of a measure are replicable over time and between different people. Previous research has shown good reliability for the STATIC-99 score and preliminary support for stable-2000 reliability. However, most of this research has been based on coding by well-trained researchers rather than frontline professionals, which may result in greater inter-rater reliability than would be found in an applied environment. To date, no research on the reliability of the inter-rater has been conducted using stable 2007.Method: This study is designed to establish the inter-rater reliability of the Static-99 and STABLE 2007 marked independently by correctional program officials trained at a sexual offender assessment unit within Canada's correctional services. Results: Overall, inter-rater reliability for U.S.-99 and STABLE-2007 field coding was excellent for total scores and ranged from acceptable to excellent for all individual elements of both measures. Conclusions: Establishing good reliability between rates on evaluation measures has the beneficial effect of ensuring accurate and high-quality assessments and supports the credibility of evaluations when reviewed by external decision-making bodies. Keywords: Static-99, Stable-2007, interrater reliability, field trialThe potential danger to public safety of those already known to have committed a sexual offence is a major concern for courts, police organizations and release institutions such as prisons and prisons. As a result, risk assessment has become a critical part of many legal decisions (e.g. civil engagement assessments, end-of-sentence assessments and treatment assignment) and is considered important information for decision-makers involved in the management of sexual offenders both by releasing organisations (e.g. prisons and prisons) and of the community. Risk assessment and risk measures have evolved considerably over the past two decades (e.g. Harris & Hanson, 2010; Hanson, 2005; Mann, Hanson, & Thornton, 2010). As part of this evolution, separate approaches and several generations of risk assessment have been identified and discussed & Bonta, 2010; Bonta, 1996; Heilbrun, 1997). Actuarial risk assessment measures (static or dynamic) help identify which offenders are most at risk of reoffending, and dynamic risk assessment measures add information related to areas of concern, potential treatment objectives, and provide attention to oversight efforts. Actuarial measures of the STATIC family including static-99 (Hanson & Thornton, 2000) and its dynamic risk measurement, STABLE-2007 (Hanson, Harris, Scott, & Helmus, 2007) are two of the most widely used tools for assessing static and dynamic risk among sex offenders in Canada and the United States (McGrath, Cumming, Burchard, Zeoli, & Ellerby, 2010). Given the seriousness of the decisions informed by the risk assessment, it is essential that these measures are considered to be: 1) reasonably standardized; 2) valid; and 3) reliable. Complete coding manuals for both static-99 (Harris, Phenix, Hanson, & Thornton, 2003) and Stable-2007 (Fernandez, Harris, Hanson, & Sparks, 2014) help improve standardization. Various studies have shown the validity of Static-99 (see www.static99.org for an overview of over 60 replicas) and STABLE-2007 (Hanson, Harris, Scott, & Helmus, 2007) to predict sexual relapse. Static-99 reliability research has consistently shown high intraclass correlations (ICC) for the total score from .90 to .98 (Barbaree, Seto, Langton & Peacock, 2001; Rettenberger, Matthes, Boer, & Eher, 2009) and ICC acceptable for individual static-99 articles (Harris et al., 2003). However, most of these studies were based on the measure being marked under research conditions, which often results in higher reliability estimates. Research studies often present better estimates of reliability, as rates generally have the same level of experience and training and obtain the scale based on identical information and conditions (e.g. time constraints). Coding risk measures into applied settings can reasonably result in lower levels of reliability. In clinical contexts risk measures are usually supplemented by frontline staff who often have limited formal training in psychometry or risk assessment. In addition, rates may have access to different information, especially if raters are conducting their interview. Particularly for some scales such as stable-2007 or PCL-R, different rates can ask different questions in an interview, and consequently get significantly different information. It is possible that the responses of offenders may also be influenced by the interpersonal style of the or the offender's mood or time of day. Even with the same accessible sources of information, real-world rates can selectively deal with different information (e.g. for example, some rates can search directly for or notes in the system, while another can assume that anything noteworthy will be mentioned in other reports). As a result, real-world reliability studies of risk scales can combine inter-rater reliability (the extent to which two raters can agree on a score) with some test-retest reliability characteristics (the extent to which the offender can provide similar information at two different times over time). However, although the concept becomes a little broader, it is still imperative to understand (as a result of all these features), the extent to which we can expect consistency in scores between different rates in real contexts. So far, three studies have examined interrater reliability for static-99 in field applications. (2012) found high levels of interrater reliability in correctional settings in Texas and New Jersey, but noted several total scores in 45% of cases. (2009) found reasonable inter-rater reliability between static-99 scores of professionals working for the signer and the respondent (ICC = .64) in SVP proceedings, but commented that the differences in score varied in the direction one would expect based on the role of the evaluator (for example, whether they represented the State or the offender). Very interesting is that Quesada, Calkins and Jeglic (2013) compared field rater scores on Static-99 with scores derived from researchers. The interrater agreement for the total score was excellent (ICC = .92) and varied from acceptable to excellent for individual items (from .62 to .94). An earlier version of the STABLE tool, the STABLE-2000 demonstrated good reliability (ICC = .89) for total scores (Hanson et al., 2007). To date I am not aware of published studies on the interrater reliability of STABLE-2007. This research study is designed to establish the inter-rater reliability level of Static-99 and STABLE 2007 at a sexual offender assessment unit within Canada's correctional services to ensure that the score of these measures in an applied environment meets acceptable levels of internal reliability. Participants in the method Participants were 55 adult males who had been convicted of a sexual offense under canada's criminal code and sentenced to a federal sentence (two years or more). Offenders serving sentences of two years or more are under the jurisdiction of the Correctional Services of Canada (CSC). During the study period (August 1, 2007 to August 1, 2008), in the Ontario region of the CSC all male federal offenders were initially housed in an evaluation unit, where they were subjected to various assessments to determine the most appropriate level of safety and the of the program. Sex offenders were offered the opportunity to participate in a specialised assessment of their sexual behavior, which was used to assess the risk of sexual recurrence and to make appropriate recommendations on the intensity of treatment. Treatment, the evaluations were completed by the correctional programme managers. In this study, sex offenders subjected to a specialised assessment of sex offenders were asked whether they would be willing to participate in a study to examine whether the measures used in the assessment were repeatable if evaluated by a different evaluator for the same individual. Only participants who agreed to participate and signed the study consent form were interviewed for the study. Procedure Participants were asked whether they would be willing to participate in a study examining the interrater reliability of the measures used by correctional program officials as part of specialized assessments of sexual recidivist offenders. It was explained that a second correctional program manager, who was blind to the Static-99 and STABLE-2007 scores completed by the first Correctional Program Manager, would complete a second interview and review of participants' institutional files and then mark Static-99 and STABLE 2007 based on their independent file review and interview information for each topic. Participants were informed that the measurement study score would not be included in their official evaluation report. Correctional program managers alternated the order of the study interview with the official evaluation interview for each subject to mitigate the effects of the order. On average, the interviews lasted about 2 hours. File information available to correctional program officers included police reports for both sexual and non-sexual allegations and convictions, documents from any provincial incarceration at a provincial correctional facility, court documents including sentencing grounds and any victims impact statement, and previous psychiatric and/or psychological reports, and previous CSC documentation. All offenders had official criminal records, relevant police reports, and previous CSC records if they ever served a federal sentence. Most participants (> 90%) had additional side information from interviews with family members. Six correctional program officers were part of the reliability study. Each had a degree in psychology or a university degree in evaluation and behavioral intervention and was trained on Static-99 and STABLE-2007 by a Static-99/STABLE-2007 certified trainer. All but one evaluator had at least one year's experience in the unit, meaning they would get about 40 Static-99s and 40 STABLE-2007s to participate in the study. An evaluator had less experience in the field, but was part of the scale development team. Each Program Officer rotated to take on the role of the second evaluator for the reliability study. Correctional program officers were blind to each other's scores. STATIC-99 measurements. The STATIC-99 is a tool designed to help predict recidivism for sexual offenders. Hanson and Thornton (1999) developed this measure based on follow-up studies from Canada and the UK with a total sample size of 1,301 sex offenders. STATIC-99 consists of 10 items and produces estimates of future risk based on the number of risk factors present in each individual. Risk factors include previous sexual offenses, current non-sexual violence, a history of non-sexual assault, number of previous sentencing dates, under the age of 25, having male victims, never living with a lover for two continuous years, history of non-contact sexual offenses, having unrelated victims, and having unknown victims. STABLE 2007 was developed to assess the change in the state of risk in the medium term, assessment needs and help predict recidivism in sexual offenders. Hanson and Harris (2000; Hanson et al., 2007) developed this risk assessment tool based on extensive prospective study from Canada, the states of Alaska and Iowa with a total sample size of 997 sex offenders. STABLE 2007 consists of 13 items and produces stable dynamic risk estimates based on the number of stable dynamic risk factors present in each individual. The risk factors included are the presence of significant social influences, the ability to stability of relationships, emotional identification with children, hostility towards women, general social rejection, lack of concern for others, impulsivity, poor problem solving skills, negative emotionality, sexual desire and concern, sex as coping, deviant sexual preference and cooperation with supervision. Analysis Overview Planned analyses were intraclass correlations (ICC) and percentage agreement of identical scores completed by the two independent correctional program managers for the 55 subjects on both individual articles and total scores for static-99 and STABLE-2007. ICC uni-way random effects were used because raters differed between cases (note that in uni-way random effect models, there is no separate ICC for consistency with absolute models). The ICC has been reported for individual measures (the average measures refer to cases where the risk score would be estimated as an average of several rates, which is generally not the way in which these scales are used or reported in practice). For dichotomous articles on Static-99, kappa is traditionally reported as the inter-rater reliability statistic, but does not take into account the variation in rates. CIC can be used for dichotomous variables (Kappas have been calculated for comparison and were always within .01 from the ICC; for example, .88 versus .89). For STABLE-2007, since the coding rules are a bit more subjective and there is a much wider range of total scores, the percentage agreement is presented for exact total scores, as well as for the agreement within 1 and 2 points. Following Cicchetti Cicchetti ICC values above .75 were considered excellent. ICC values between .60 and .74 were considered good and between .40 and .59 were considered fair. Results Table 1: Intraclass correlations for total static-99 score Young people ever experienced with non-sexual violence index Preventive non-sexual violence Previous sexual offences Non-contact dates Unrelated victim Foreign victim Male victim Table 2: Intraclass correlations for total score STABLE 2007 72 (within 2 Points) 48 (within 1 point) 13 (identical score) Significant social influences Ability to relate emotional ID with children Hostility to women General social rejection Lack of concern for others Impulsive Poor problem solve negative emotionality Sexual desire/concern Sex as coping deviant sexual preferences Cooperation with supervision Note. Due to the fact that a badger left some elements empty, the sample size was 53 cases for general social rejection, sex such as coping and sexual guidance/concern. For emotional identification with children, only 39 cases were analyzed, as this article is not rated for offenders with non-victims under the age of 14. For this element, in three cases, one rate marked the element and another omitted the assessment; analyses are presented only when both have marked the element. As can be seen from Tables 1 and 2, the agreement was raised between the two independent raters for the total scores on both STATIC-99 (.95) and STABLE-2007 (.86). The total scores were identical between the two raters for 78% of cases in the Static-99 and were within 2 points for 72% of stable-2007 cases (scores were within 1 point for 48% of cases and identical in 13% of cases). CIC for individual STATIC-99 elements ranged from .57 to .93 with a median of .88. Scores were identical between both rates over 90% of the time out of 7 of the 10 individual items. The two STATIC-99 elements with the lowest interrater reliability were contactless offences (ICC = .71) and the Non-Sexual Violence Index (ICC = .57); all other articles had an ICC in the excellent range. ICT for individual STABLE-2007 items ranged from .37 to .83 with a median of .70. Four pics of items were in the excellent range, six were good, two were rater (ICC for poor troubleshooting = .55, Negative emotion ICC = .51), and an ICC fell just below fair (Significant social influences ICC = .37). Despite the more modest ICT for STABLE items, the percentage agreement was still generally good (although lower than for Static-99 items). Nine of the 13 entries had identical scores more than 70% of the time (median percentage agreement = 72%). The of relationship stability had the highest percentage agreement with identical scores among evaluators 91% of the time. The lowest agreement concerns significant social influences (61%). Overall, inter-rater reliability for field coding of statics-99 and STABLE-2007 has been for total scores and generally varied from acceptable to excellent for individual items, with some STABLE items falling in fair regions and below fairs. Discussion Establishing good interrater reliability on assessment measures has the beneficial effect of ensuring accurate and high-quality assessments of sexual offenders. In addition, evidence of high internal reliability provides support for the credibility of completed assessments when examined by external bodies, such as the National Parole Board or courts during hearings for dangerous offenders or civil engagement hearings. This study provides proof of the interrater reliability of the Static-99 and STABLE 2007 completed in an environment applied by correctional program managers. The high levels of agreement for the total scores for both the Static-99 and the STABLE 2007 and for many individual items for both measures are remarkable. Of course, inter-rater reliability has fallen below the voucher threshold for a STABLE-2007 entries. Generally, the reliability of individual articles on STABLE-2007 was more modest than static-99, suggesting greater variability between raters. This is not surprising given that these elements require more discretion and subjectivity in evaluation than in Static-99 articles (and also greater dependence on interviewing, which differed between evaluators). Despite the lower reliability of some of these individual articles, however, the total scores on STABLE-2007 still demonstrated excellent reliability. By way of illustration, the two points of both measures with the lowest levels of inter-mediterranean agreement will be discussed in more detail. On Static-99, contactless sexual offences and the non-sexual violence index had the lowest level of agreement. Examination of cases where raters were uneven on these elements suggested that rates would benefit from the reference to the coding manual for static-99 during the completion of the scale. For example, it turned out that a rater believed that a conviction for Pronounced Threats did not meet the Index's non-sexual violence criteria despite Minnacioso's inclusion on the list of crimes included in the STATIC-99 coding manual as examples of convictions for non-sexual assault (pg. 28). Similarly, it turned out that the discrepancy on the article Contactless Sexual Offences was mainly due to confusion about the difference between convictions for possession of child pornography (a non-contact sexual offence) and child pornography (a contact offence). These results have important implications because in some jurisdictions Static-99 scores weigh heavily in decisions on positioning, and the management of sexual offenders. A review of total discrepant scores on Static-99 revealed that the discrepancy led to a different nominal risk category (low, moderate-low, moderate-high, high) in 3 out of 55 cases (5%), (5%), in one case, the total score of an evaluator placed the subject in the nominal category at medium-high risk, while the total score of the other evaluator placed the subject in the high-risk nominal category. In two cases, the difference between the score of the two evaluators led to differences between the medium-low risk nominal category and the medium-high risk nominal category. On STABLE 2007, the voices with the lowest interrater agreement included significant social influences and negative emotion. Both elements appeared to be heavily influenced by discrepant information provided to the interviewer during independent interviews. Participants identified significant social influences that were dissimilar to several interviewers, or identified different social influences to one interviewer and few or none to the other interviewer. Discrepancies in the negative emotionality article also appeared to be related to the discrepant information provided during interviews. It seemed that some participants presented themselves as hostile and ruminant during an interview and then denied similar thoughts and emotions during the next interview. It is possible that the demand characteristics of the two interviews were incoherent with the quality and type of information provided. A broader assessment of interview notes from both sessions may reveal whether offenders interviewed for the study will feel less obligated to present themselves in a favorable light because they are aware that the information would not be used in the official report or to make decisions about their case within Canada's correctional services. It is also possible that negative emotion is more susceptible to natural mood fluctuations, or that differences reflect reactions to different interview styles between raters. This would provide further evidence of the importance of having good side information when completing evaluations using STABLE 2007. Overall, it was found that many of the differences in stable score could be attributed to differences in information compared to differences in the interpretation/application of coding rules. As a result, studies where raters get scale based on the same information could achieve greater inter-rater reliability. This study, however, is more realistic than typical field settings where different evaluators tend to conduct their own interviews, which is likely to affect scoring decisions. A review of the discrepant total scores revealed that the discrepancy led to a different category of nominal risk/need (low, moderate, high) in 6 out of 54 cases (11%). In all but one of the cases, the between the score of the two evaluators was among the nominal categories of moderate risk/need and high/risk need (in the other case, a one-point scoring difference placed the offender in the low risk/need category by an evaluator and in the moderate/moderate category from another). Similar to quesada et al. While encouraging with regard to field coding in the United States-99 and STABLE-2007, this study is based on a small sample (n = 55) in a unit specializing in assessments of sexual offenders using these tools. Replication is required with larger sample sizes. Also, as Quesada et al. note. It should also be noted that the correctional program managers involved in the ongoing study had all been trained by a certified trainer, all but one had previously obtained at least 40 Static-99 and STABLE-2007, and were under the direct supervision of a fully experienced psychologist who was a certified trainer (although the evaluations used for this study were before the cases were discussed with the supervisor). This may explain the greater reliability of that found by Murrie et al. This speaks to the importance of safeguarding quality by ensuring that evaluators are properly trained, have clinical supervision by an experienced evaluator, and preferably have a peer review or mentoring system to help address potential rate drift. Hanson, Helmus and Harris (2015) found that the Conscientious Evaluators of the Dynamic Supervision Project (i.e., those who filled out and submitted all forms) were more accurate (AUC in the range 0.76 to 0.80 on the various versions of STABLE and Static-99 = 343) than the assessments presented by the less conscientious evaluators (AUC in the range 0.58 to 0.68; n = 421). As a result, when paying attention to the training, supervision and quality of assessments, and evaluators are engaged in the process, risk assessment tools are both more reliable and potentially more accurate in predicting sexual reoffending, both of which are critical to maximising public safety when decisions are based on the results of actuarial and dynamic risk assessment tools, while increasing the fairness and consistency of assessments with profound implications for the civil liberties of offenders. ReferencesAndrews, D. A., & Bonta, J. (2010). The psychology of criminal conduct. Cincinnati: Anderson Publishing Co. Barbaree, H., Seto, M., Langton, C., & Peacock, E. (2001). Assessment of predictive accuracy of six risk assessment tools for offenders. Criminal justice and conduct, 28, 490-521. doi: 10.1177/009385480102800406Boccacini, M. T., Murrie, D.C., Mercado, C., Quesada, S., Hawes, S., Rice, A. K., & Jeglic, E. E. (2012). Implications of static-99 field reliability results for using and reporting scores. Criminal justice and 39, 42-58. doi: 10.1177/0093854811427131Bonta, J. (1996). Assessment and treatment of risk needs. In A. T. Harland (Ed.), Choosing correctional options that work: Defining the demand and evaluating the supply (pp. 18-32). Thousand Oaks, CA: Sage.Cicchetti, D. V. (1994). Guidelines, criteria and empirical rules for the evaluation of standardized and standardized evaluation tools in psychology. Psychological evaluation, 6, 284-290. (2014) Fernandez, Y., Harris, A. J. R., Hanson, R. K., & Sparks, J. STABLE-2007 CODING MANUAL - revised in 2014. Unpublished report. Ottawa, ON: Public Safety Canada.Hanson, R. K. (2005). Twenty years of progress in assessing the risk of violence. Journal of Interpersonal Violence, 20, 212-217. doi:10.1177/0886260504267740Hanson, R. K., & Harris, A. J. R. (2000). The sexual harasser needs rating assessment (SONAR): A method of measuring change in risk levels. User report 2000-01. epartment of the Attorney General of Canada, Ottawa. Available from www.ps-sp.gc.ca/res/cor/rep. Hanson, R. K., Harris, A.J.R., Scott T.L. & Helmus, L. (2007) Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision project. Ottawa: Public Safety, Canada.Hanson, R. K., Helmus, L.M., & Harris, A. J. R. (2015). Assessment of the risk and needs of supervised sex offenders: A prospective study using STABLE-2007 and Static-2002R. Criminal justice and behaviour, 42, 1205-1224. doi:10.1177/0093854815602094Hanson, R. K., & Thornton, D. (2000). Improving risk assessments for sex offenders: a comparison of three actuarial scales. Human law and behavior, 24, 119-136. doi:10.1023/A:1005462921333Hanson, R. K., & Thornton, D. (1999). STATIC-99: Improve actuarial risk assessments for sexual offenders. User report 1999-02. Department of the Attorney General of Canada, Ottawa, Canada. Available from www.ps-sp.gc.ca/res/cor/rep.Harris, A. J. R., & Hanson, R. K. (2010). Clinical, actuarial and dynamic risk assessment of sexual offenders: why do things keep changing? Sexual assault journal, 16, 296-310. doi:10.1080/13552600.2010.494772Harris, A., Phenix, A., Hanson, R. K., & Thornton, D. (2003). Static-99: Coding rules revised in 2003. Ottawa, ON: Solicitor General Canada.Heilbrun, K. (1997). Management forecasting models relevant to risk assessment: the importance of the legal decision-making environment. Human law and behavior, 21, 347-359. doi:10.1023/A:1024851017947Mann, R. E., Hanson, R. K., & Thornton, D. (2010). Risk assessment for sexual relapse: Some proposals on the nature of psychologically significant risk factors. Sexual abuse: A Journal of and Treatment, 22, 191-217. doi:10.1177/1079063210366039McGrath, R. J., Cumming, G. F., Burchard, B. L., Zeoli, S., & Ellerby, E. (2010). Current practices and emerging trends in sexual abuse management: The Safer Society 2009 North American Survey. Brandon, VT: Safer Company D., Boccacini, M., Turner, D., Meeks, M., Woods, C., & Tussey, C. (2009). Badger (dis)agreement on risk assessment measures in sexually violent predatory proceedings: evidence of contradictory fidelity in forensic assessment? Psychology, Public Order and Jurisprudence, 15, 19-53.doi:10.1037/a0014897Quesada, S. P., Calkins, C., & Jeglic, E. L. (2013). An examination of the reliability of the interrater between professionals and researchers on static-99. International Journal of Offender Therapy and Comparative Criminology, 58, 1364-1375. doi:10.1177/0306624X13495504Rettenberger, M., Matthes, A., Boer, D., & Eher, R. (2009). Prospective actuarial risk assessment: a comparison of five risk assessment tools in different subtypes of sexual offenders. International Journal of Offender Therapy and Comparative Criminology, 2, 169-186. doi:10.1177/0306624X08328755 Footnote 1The static family of scales collectively refers to Static-99, Static-99R, Static-2002, and Static-2002R. Address by Yolanda M. FernandezCorrectional Services of CanadaRegional Headquarters (Ontario)433 Union St.Kingston, OntarioK7L 4Y8Yolanda.fernandez@csc-ccc.gc.ca © 2006-2020 Pabst Science Publishers Publishers