Safety in Numbers

303 Bagot St. Ste. 306
Kingston, Ontario
K7K 5W7
Canada
1 855 955 3282

Static-99R & Static-2002R/BARR-2002R Training Guidelines

January 11, 2021

For information concerning these guidelines, contact Dr. Yolanda Fernandez at Yolanda.Fernandez@saarna.org or toll free at 1 855 955 3282

These training guidelines were drafted by a collaborative working group of professionals experienced in sexual offender risk assessment, with input from the SAARNA Board of Directors. The SAARNA Board of Directors would like to thank the following members of the working committee for their valuable input and dedication to advancing the use of actuarial measures as part of quality risk and needs assessment.

Dr. Craig Teofilo (lead)
Dr. Dana Anderson
Mr. Jacob Bezanson
Dr. Dawn Pflugradt
Mr. Robert Colley
Dr. Deirdre D'Orazio
Dr. Erik Fox
Dr. Maaike Helmus
Mr. Corey McNally
Dr. Martin Rettenberger
Ms. Megan Stenswick
Dr. Wineke Smid
Dr. Pamela Yates

This document is directed toward certified trainers of the Static measures and the administrators of the jurisdictions in which these measures are routinely used. It provides training guidelines for Static-99R, Static-2002R and BARR-2002R risk tools (the Static measures). In brief, the recommendations are as follows:

A. **Number of participants**. The recommended number of participants is up to 25 for one trainer and up to 50 for two trainers. We strongly recommend against providing trainings with more than 40 participants (one trainer) or 60 participants (two trainers).
B. The **duration of the training** should be between eight hours and 12 hours.
C. **Retraining**. Individuals should be retrained every two to five years after their previous training. We recommend that individuals do not use Static measures if they have not received training within the past 10 years.
D. **Certification.** We recommend that trainees be given the option to receive a competency certificate based on passing a test or exercise. Minimally, trainees should receive a certificate of attendance.

Risk Assessment: General Framework and Professional Responsibilities

Risk assessment is one of the most ubiquitous and important activities in the criminal justice system. It affects (or should affect) virtually every decision made about individuals charged or convicted of offences, including bail, sentencing, security classification, treatment referrals/intensity, parole, and community supervision intensity and conditions. Risk assessment balances public safety and the rights of individuals charged or convicted of offences. Done appropriately, evidence-based risk assessment improves transparency and accuracy of criminal justice decision-making and allocates limited resources to maximize effectiveness and public safety.

Although risk assessment is a pillar of good forensic practice, it is important to be aware of key ethical or professional guidelines and responsibilities for the appropriate use of risk assessment tools (e.g., American Psychological Association's Forensic Specialty Guidelines, 2013[1]; Association for the Treatment of Sexual Abusers' Adult Practice Guidelines, 2014). These guidelines may vary by jurisdiction or profession, but generally, the use of evidence-based risk assessment tools requires appropriate training in how to score and interpret the scale and knowledge of the scientific foundation for the approach used. This would include risk factors for sexual and general recidivism and the empirical strengths and limitations of the scale (e.g., reliability/validity overall, as well as for subgroups of offenders). However, jurisdictions vary in education and experience requirements for conducting risk assessments. As such, there is variation in expertise and familiarity with the empirical background research among users of risk

[1] Although focused towards psychologists, these guidelines are meant to be informative to diverse professionals involved in applying psychological research in working with individuals in the criminal justice system.

assessment scales. Some knowledge of these issues and guidelines is required to make decisions about risk scale selection and how that risk scale would relate to decision-making in an individual jurisdiction.)

Static-99R

Static-99R is intended to rank offenders in terms of their relative degree of risk for sexual recidivism based on commonly available demographic and criminal history information that has been found to correlate with sexual recidivism in adult male sex offenders. When combined with an appropriate table of norms (e.g., Phenix et al., 2016), Static-99R characterizes the individual's relative risk for sexual recidivism in terms of how unusual it is (using percentiles) and how it compares to the risk presented by the typical sex offender (using risk ratios). Available norms also describe estimates of absolute recidivism rates over fixed follow-up periods. The information provided by Static-99R can be thought of as a baseline estimate of risk for new sexual charges and convictions. This baseline assessment can be used to guide treatment and supervision strategies designed to reduce the risk of sexual recidivism.

Static-2002R/BARR-2002R

Static-2002R is similar in many ways to Static-99R; it is intended to rank offenders in terms of their relative degree of risk for sexual recidivism based on commonly available demographic and criminal history information that has been found to correlate with sexual recidivism in adult male sex offenders. Compared to Static-99R, it was designed to have greater conceptual clarity and to improve case conceptualization and management. Both scales predict sexual recidivism with similar accuracy, but they also add incrementally to each other (Babchishin et al., 2012). Similar normative data are available for Static-2002R as per Static-99R (e.g., Phenix et al., 2016). BARR-2002R consists of a subset of Static-2002R items summed separately; they assess risk of general and violent recidivism for individuals with a history of sexual offending.

Training Guidelines

A growing body of evidence suggests the quality of training is related to predictive accuracy (Hanson et al, 2015). California probation and parole officers demonstrated remarkably high predictive accuracy and inter-rater reliability (Hanson et al, 2014). This is likely the result of a California statute mandating those who rate registered sex offenders on probation and parole be trained by a certified Static trainer and to attend a recertification training every two years (as well as other quality control systems in place). In an upcoming meta-analysis of field validity studies of Static-99R, predictive accuracy is significantly higher for studies known to have used a certified trainer compared to studies with unknown training practices (Helmus et al., 2020).

Ideally, practice standards should be based on empirical evidence, which is unfortunately not currently available for the Static measures. Consequently, the present guidelines are a collection of recommendations based on the professional opinions of experienced individuals who are committed to the process of high-quality risk assessment. The present guidelines specify best practices, a range of acceptable practices, and recommendations not to go beyond certain limits based on the collective experiences of Trainers and Master Trainers in the Static measures. The guidelines encompass a similar spirit to the user guidance currently provided in the coding manuals for other structured risk prediction measures. Individual jurisdictions or organizations are encouraged to develop their own standards congruent with these guidelines. For example, we strongly encourage that jurisdictions or organizations set and enforce a reasonable timeline standard for user recertification. Adherence to the standards set by a jurisdiction or organization is the responsibility of that group.

Compiling Best Practices

Training takes place in various contexts and for different purposes;therefore, there is an inherent need for flexibility to modify the training materials. However, there must still be a minimum standard in order to maintain the reliability and validity of the measure and to ensure its accuracy in decisions that affect peoples' lives.

These training guidelines for the Static measures were developed by a working group of certified trainers with input from the SAARNA Board of Directors. The working group included veteran certified trainers from multiple jurisdictions across North America, Europe, and New Zealand. The project of compiling best practices for Static training began with collecting and reviewing a survey of recommendations and observations of certified Static trainers.

The training guidelines for Static-99R and Static-2002R/BARR-2002R training are offered to assist in ensuring accurate and reliable scoring and interpretation of the measure. The standards apply to in-person and web-based live trainings[2].

Training guidelines are provided on the following issues:
1) number of training session participants;
2) length of training sessions;
3) frequency of retraining; and
4) provision of certificates (including certification tests).

[2] At this time, we do not support on-demand training sessions (i.e., watching a videotaped training) as these do not allow for interactions between the certified trainer and attendees.

1. Number of Participants

- **Overview:**

The context of training sessions can vary widely. Participants in group training sessions may include any combination of new scorers who have never used the measure, veteran scorers who use the measure daily or weekly, and scorers who frequently testify in court and who therefore have a stronger interest in research related to the measures' development and predictive accuracy. Participants may also include administrators, policy experts, legal professionals, and others who will never score the measure, but who nonetheless need to be familiar with the measure.

Similarly, both the trainer's experience and the circumstances under which trainers deliver training sessions vary. For example, some certified trainers conduct trainings infrequently while other trainers conduct training regularly. Some certified trainers provide training to only one jurisdiction, whereas others provide training sessions in multiple jurisdictions. Some trainers always provide training sessions alone, while others train with a co-facilitator.

These and other setting factors, either individually or in combination, can influence the effectiveness of a given training session and, therefore, the number of participants that is manageable for that session.

- **Guidelines:**

Single Trainer:

Best practice: The number of participants should be no more than 25 participants for training sessions delivered by a single trainer.

Range of acceptable practices: The number of participants may be expanded up to 40 participants depending on the experience, skills, and comfort of the trainer (i.e., more experienced trainers may feel more comfortable with more participants).

Limit: More than 40 participants in a training session with a single trainer is not recommended.

Co-facilitated Training:

Best practice: When training sessions are conducted by co-trainers, the number of participants should be no more than 40 participants.

Range of acceptable practices: The number of participants may be expanded up to 60 participants depending on the experience, skills and comfort of the co-trainers (i.e., more experienced trainers may feel more comfortable with more participants

*Limi*t: More than 60 participants in a training session with co-facilitators is not recommended.

2. Duration of Training Sessions

- **Overview:**

Training sessions vary in duration and experience level of participants. Training sessions need to be long enough to allow attendees sufficient learning and practice time, including scoring exercises on multiple vignettes, and to allow discussions/questions with the trainer(s). It is also important to note that the 2016 coding manual is considerably more complex than the previous version (2003), and so may require longer training sessions than previously.

- **Guideline:**

Best practice: An initial Static training session is suggested to be between 8 and 12 hours (including a one-hour lunch and 15-minute morning/afternoon breaks) based on the needs of the participants. When training novice audiences with no experience or limited experience working with individuals who have committed sexual offences, we recommend the trainer include more extensive background information. The nature of this background information would vary based on the needs of the participants/jurisdictions, but should include research information on offender characteristics, sexual recidivism rates, actuarial risk assessment, treatment practices/effectiveness, methods of managing risk, and links to associated treatment/programs. Note that some introductory information is necessary in all training sessions (per the sample training slides), but the exact topics and depth may be tailored to different contexts. This introductory information for novices would be outside the recommended training time and may be a prerequisite training component before completing the Static training.

Range of acceptable practices: In all trainings, but particularly online, we consistently receive feedback that there is a lot of information to process about the scale and participants tend to lose focus at the end of the day. Trainers should consider holding the training over two days (e.g., two 5 or 6-hour training days is more manageable than a 9-hour training day with shorter breaks).

Limit: The minimum goal of training should include acceptable scoring in most cases, and awareness that there are additional rules for rare, complex cases. This can be done in an 8-hour training session. For professionals with experience in the areas of risk assessment and working with sexual offenders an 8-hour training day may be possible. Fewer than 8 hours is not recommended for training on either Static-99R or Static-2002R/BARR-2002R.

- **Considerations:**

The time spent in training should be considered alongside length of breaks. If delivering a live training where longer breaks are expected (e.g., people need to go off-site to purchase lunch), it may be impractical to schedule a training in the shorter end of the range above.

Note that even with a 12-hour training session, it is impossible to review every rule in the coding manual. Participants should be informed that participation in a live (either face-to-face or on-line) training is one component of getting ready to score the measure, along with a more in-depth reading of the coding manual.

3. Frequency of Retraining

- **Overview:**

The goal of training is to ensure participants can accurately and reliably score the measures. Participating in retraining on a regular basis helps to refresh constructs for participants, allows the opportunity to update scorers with new research or scoring changes, reduces coding drift, and can mitigate the misapplication of coding rules. A misinterpreted coding rule can become embedded in an organization and over time nearly everyone in the organization would score a particular item consistently, but incorrectly.

The need for retraining may be recommended by a quality assurance/review process for reasons such as repeated scoring errors even after coaching or a break in scoring for a prolonged period (e.g., 6+ months).

- **Guideline:**

Best practice: Retraining sessions are recommended every two to five years based on the needs of the individuals scoring the measures.

Range of acceptable practices: More frequent retraining may be required if users are not required to score the measures on a frequent basis. In some cases, scorers may go a period of time without using Static measures, for example, due to job changes or temporary leave. For individuals who have not scored Static measures for a period of two years, recommendations depend on the individual's levels of experience. It may be a (brief) refresher training is sufficient, or it may be that a (full) retraining is needed.

- An individual who was trained on the scale and who does not administer it within two years of training, should be retrained in the measure before scoring the scale in practice.

- An individual who was trained and scored the scale regularly (e.g., several cases a month or more), should seek consultation with a veteran scorer or certified trainer to determine whether re-training or refresher training is needed prior to scoring the scale in practice.

- An individual who was trained and was a moderate or occasional user of the scale, a refresher training may be ideal. They should be encouraged to discuss options with a certified trainer. The length of a refresher trainer should be based on an assessment by

the certified trainer of the needs of the individual or group requiring refresher training. A minimum half day is recommended.

- In jurisdictions with formal review of scoring, an individual who was trained and was a moderate or occasional user of the scale may only require intensive supervision. They should be encouraged to discuss options with a certified trainer.

Limit: More than 10 years without retraining on any measure is not recommended.

- **Considerations:**

Occasionally the Static Development Team makes recommendations for all users to obtain new training, re-training, or refresher training (e.g., after a new coding manual is produced), as appropriate given the nature and extent of the changes.

We recognize the need for additional translations for the Static measures' coding manuals and training materials for non-English speaking users. Some jurisdictions or countries may not have the capacity or infrastructure for retraining based on the best practice timeframe. In these instances, the ideal timeframe should be considered a goal to strive for. It may take some time to develop standards based on these training guidelines.

4. Certificates

- **Overview:**

Certificates help attendees with record-keeping and training dates. Certificates provide attendees with written verification of being trained by a certified trainer. In training that requires taking and passing a test vignette, the certificate conveys some level of mastery scoring the measure. Finally, certificates may be a requirement in some jurisdictions.

- **Guidelines:**

Best practice: Certification - This attests that the person has attended the training sessions and passed some type of testing/certification standard suggesting reasonable mastery in scoring the measure.

This is more common and strongly encouraged in jurisdictions that mandate the use of the Static measures, where all staff are required to be trained and where there is a policy in place to provide trainings on an ongoing basis.

Range of acceptable practices: Attendance certificates - This certificate attests that the individual has attended the training session but is not an indicator that they have met any particular learning standard(s).

These certificates are more common in one-time training sessions where it may not be practical to incorporate a post-training test (e.g., large pre-conference workshops, training sessions organized for diverse audiences, such as an event hosted by a university).

Limit: Participation without some formal recognition in the form of a certificate confirming participation in the training session is not recommended.

- **Considerations:**

It is preferred that participants successfully complete some type of test at the end of training sessions. One or more case vignettes are particularly encouraged, but other test formats (e.g., multiple choice or short answer questions) can also be incorporated. Note that you would rarely need to see perfect scores to pass a test, but a requirement for passing should include a reasonably small number of errors. Interrater reliability standards of +/- 1 point of scoring is a good bar to consider if someone should pass the certification case provided during the training.

Testing has several advantages, which is why we encourage trainers to incorporate them. They provide added assurance of competent scoring, which is necessary for good-quality implementation of risk scales. The awareness of being tested at the end of the training could promote enhanced participant engagement. Should a participant not successfully complete the test, the course of action chosen may be to review participants' responses with them, to administer a make-up test, to suggest remedial action, and/or advise that the participant should not score Static measures.

Additional suggestions/tips from various jurisdictions:

- New South Wales and the Washington Department of Corrections have certified trainers who manage a special email address for Static-99R questions from their respective staffs. This allows them to manage many commonly asked, straightforward questions. They can also vet the more complicated questions that might be sent to the general Staticquestions@gmail.com account.

- New Zealand Department of Correction's quality assurance process requires every risk assessment measure completed by a psychologist be reviewed by a clinical supervisor with a separate reviewer (usually a psychological manager or senior psychologist) involved in the review of the psychological report that conveys the results. In addition, the application of risk assessment measures is supported by Subject Matter Experts (SME's) who provide guidance, advice and review of scoring and interpretation if it is sought by the probation officer or psychologist.

- Certified trainers can create their own Frequently Asked Questions (FAQs), particularly jurisdiction specific.

- Trainers who deliver ongoing training sessions in a particular jurisdiction could create a document that is supplementary to the Static-99R and Static-2002R coding manuals that tracks guidance they have given about how to code particularly unique situations in their jurisdiction that may not be addressed in the scoring manual (note that this suggestion increases consistency but does not guarantee accuracy. Therefore, we recommend this be shared with the Static Development Team of SAARNA).

- For trainers: Training materials are available on the SAARNA.org website within the Trainers section, including training slides, sample vignettes, test vignettes, and sample certificates (attendance and certification).

- For users and jurisdictions: When you have questions about a specific training situation, please contact a certified trainer. If you are unable to contact a certified trainer send your questions to the Staticquestions@gmail.com account.

References

American Psychological Association (2013). Specialty guidelines for forensic psychology. American Psychologist, 68, 7-19. doi: 10.1037/a0029889

Association for the Treatment of Sexual Abusers (2014). *ATSA practice guidelines for assessment, treatment interventions, and management strategies for male adult sexual abusers.* Beaverton, OR: Professional Issues Committee, ATSA.

Babchishin, K. M., Hanson, R. K., & Helmus, L. (2012). Even highly correlated measures can add incrementally to actuarial risk prediction. *Assessment*, 19, 442-461. doi:10.1177/1073191112458312

Hanson, R. K., Phenix, A., Lunetta, A., Neeley, J. & Epperson, D. (2014). The field validity of Static-99/R sex offender risk assessment tool in California. *Journal of Threat Assessment and Management, 1*, 102-117. doi:10.1037/tam0000014

Hanson, R. K., Helmus, L. & Harris, A. J. R. (2015). Assessing the risk and needs of supervised sexual offenders: A prospective tudy using STABLE-2007, Static-99R, and Static-2002R. *Criminal Justice and Behavior, 42*, 1205-1224. doi: 10.1177/0093854815602094

Helmus, L. M., Hanson, R. K., Murrie, D. C., & Zabarauckas, C. L. (2020). *Field validity of Static-99R and STABLE-2007 with 4,433 men serving sentences for sexual offences in British Columbia: New findings and cumulative meta-analysis.* Manuscript under review.

Phenix, A., Helmus, L., & Hanson, R. K. (2016). Static-99R & Static-2002R evaluators' workbook. Available at www.static99.org