

## **A Correctional Agency's Consumer Guide to the Selection of a Risk Assessment Instrument**

©Ralph C. Serin  
Carleton University, Ottawa, Ontario Canada

Ratings consider Capacity, Implementation Issues, and Empirical Evidence. A “passing” grade for all areas is required in order for a positive decision to use or pilot the proposed risk assessment scale.

### **A. Capacity:**

1. Does your agency have the resources to use the proposed scale?
  - a. Financial – can you afford it in terms of initial/implementation and long-term operating costs?
  - b. Staff – do you have sufficient staff with the appropriate background to implement?
2. Does the proposed scale answer the question(s) of interest?
  - a. Custody classification
  - b. Generic risk assessment
  - c. Pre-release
  - d. Programming
3. Does it fit your policy and legal framework?
4. Does it fit your business plan and overall mission?
5. Does it advance your correctional agenda?
6. Can you explain it to legislators?
7. Will it complement or replace existing instruments (if replace, will a comparison be made during the pilot of efficiency or effectiveness)?
8. Do you need a static scale, a dynamic risk scale, or one that does both?

### **B. Implementation Issues:**

1. How long does it take to complete the instrument?
2. Is the training skills-based using case studies? Is the required training model manageable for your agency?
3. Is the necessary information already available for its completion or will new information (file/interview) be required?
4. What are the training requirements – duration, competency, inter-rater, university degree, etc.
5. Can the agency provide quality assurance and oversight?
6. Is there a cost for an accreditation process. If so, what is the cost and will it minimize liability?

## **C. Empirical Evidence<sup>1</sup>:**

### ***Predictive Validity***

1. Does it include items that predict recidivism? The Area Under the Curve should at a minimum exceed 0.70, and preferably exceed 0.80. If simple correlations are provided, then  $r > .25$  would be minimal as this is the current minimum benchmark for many risk scales. If for a unique group, are unique items included?
2. Does it capture changes in risk state over time?
3. Does it predict future re-offending (of the type of interest – general, violent, sexual)
4. Has the instrument been validated on a population similar to yours? Are there norms for cut-points that relate to different rates of success and failure?

### ***Other Issue:***

5. If self-report, does it address offender malingering?
6. Is the scale culturally and gender informed?

### ***Construct Validity***

1. Does the instrument reflect the accepted risk and need domains, as suggested in the literature.
2. Are their unique items (sexual deviance, mental health symptoms, breach of no-contact orders) for unique subgroups (SO, MDO, DV)?
3. Have analyses been completed regarding internal structure (do the items fit within a subscale)?
4. Have analyses been completed to demonstrate that domains relate to validated measures of a similar construct?

### ***Reliability***

1. Is there evidence of test-retest reliability (will the offender obtain similar scores over short time intervals)?
2. Is there evidence of inter-rater reliability (will an offender receive similar scores if assessed by two different staff)? If intra-class correlations are provided, it should exceed  $r = .80$  (this is preferable to a simple Pearson  $r$ ). If kappa is provided, it should exceed  $k = .65$ . If percentage agreement is provided, it should be 80% or greater. The more specifically a risk score is solely used for a *key* decision, then the threshold for reliability should be higher.

---

<sup>1</sup> See Skeem, J.L. & Eno Louden, J. (2007). Assessment of evidence on the quality of the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS).

<b>A Correctional Agency's Consumer Guide to the Selection of a Risk Assessment Instrument</b>
--

©Ralph C. Serin  
 Carleton University, Ottawa, Ontario Canada

**Rating System:** Pass reflects a score of 15 or greater.

<b>Does Fully Meets Needs</b> (3)	<b>Mainly Meets Needs</b> (2)	<b>Slightly Meets Needs</b> (1)	<b>Fails to Meet Needs</b> (0)
--------------------------------------	----------------------------------	------------------------------------	-----------------------------------

Evaluation Domain Considered	Rating
<b>Capacity</b>	
Affordable both in terms of fiscal and staff costs.	<i>3/2/1/0</i>
Answers the question(s) of interest.	<i>3/2/1/0</i>
Fits policy and legal framework.	<i>3/2/1/0</i>
<i>Evaluation</i>	Score = ___
<b>Implementation</b>	
Implementation requirements (time, cost).	<i>3/2/1/0</i>
Training requirements (build competency, staff background).	<i>3/2/1/0</i>
Quality assurance (availability of requisite information).	<i>3/2/1/0</i>
<i>Evaluation</i>	Score = ___
<b>Empirical Evidence</b>	
Predictive validity (related to re-offending).	<i>3/2/1/0</i>
Sample representativeness (norms available).	<i>3/2/1/0</i>
Construct validity (measures appropriate constructs).	<i>3/2/1/0</i>
Reliable (inter-rater > .80)	<i>3/2/1/0</i>
<i>Evaluation</i>	Score = ___
<b>Overall Evaluation</b> <b>(Total score must ≥ 15 for consideration)</b>	<b>Implement/ Pilot/Reject</b>