



**Research Brief**  
**Reliability of the Static Risk Offender Need Guide for Recidivism**  
**(STRONG-R)**

Xiaohan Mei, M.A.  
Zachary Hamilton, Ph.D.  
Washington State University



## Reliability/Internal Consistency of STRONG-R

Reliability is an essential criterion for accurate estimates of individuals' risk and needs. In order to obtain an accurate assessment, the estimation must demonstrate an acceptable level of both measurement validity and reliability. Reliability is also viewed as partial evidence of the internal structure of an assessment instruments (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 1999). One of the indispensable aspects of reliability is internal consistency. The current research in brief describes evidence of internal consistency and reliability of the Static Risk Offender Need Guide – Revised (STRONG-R) instrument.

### Method

Scale reliability (construct reliability) refers to as the internal consistency of measurement. In social and behavioral science studies, Cronbach's alpha coefficient is a generally accepted approach to examine the reliability of a scale (Cronbach, 1951; Bollen, 1989). A scale represents a composite measure, which is an unweighted sum of all items within a given scale. Cronbach's alpha is a good estimate of internal consistency when all the items are, at least, tau-equivalent (Wang & Wang, 2012). In other words, in the theoretical framework of Item Response Theory (IRT) and Classical Test Theory (CTT), all items in a scale must have statistically equivalent discriminate values (loadings); otherwise, Cronbach's alpha coefficient is considered a biased estimate of a scale's internal consistency (Raykov, 2001).

However, if a scale is neither tau-equivalent nor parallel, scale reliability can still be accurately evaluated via Omega reliability (Joreskog, 1971; Dillon & Goldstein, 1984). Omega reliability does not assume item invariance within a given scale, and it is the "*ratio of the variance explained by the true score to the variance to the observed variance of the latent construct* (Wang & Wang, 2012, p. 86)"

In the current study, Cronbach's alpha coefficients were computed for scales that were identified as tau-equivalent, and Omega reliability coefficients were computed for those that were congeneric. Higher numerical values of coefficients indicate higher internal consistency<sup>1</sup>. For first-order constructs, the Cronbach's alpha coefficients are interpreted as the total information a scale provides as functions of discrimination and difficulty, which is also referred to as Item Information Function (Thoman, 2011).

### Results

#### *First Order Construct/Scales*

As presented in Table 1, the Cronbach's alpha coefficients for first-order factor/sub-scales ranged from .631 (*Mental Health Issue - Recent*) to .151 (*Aggressiveness- Fixation*). The internal consistency of the first order factor/sub-scales is relatively low. However, according to Cortina (1993), Cronbach's alpha estimates perform less than ideal when scales have few items;

---

<sup>1</sup> It should be noted, because the psychometric analysis was employed within the framework of (IRT) or Item Factor Analysis (IFA), the interpretation of the Cronbach's alpha coefficients needs to be further addressed. Unlike Classical Test Theory (CTT), IRT or ITA assumes no measurement errors, including systematic and random error. The traditional interpretation of the Cronbach's alpha coefficient is the ratio of true score variance, or the total variance in a given scale without measurement errors. Because we utilized IRT (or ITA) and Higher Order Modeling (see Mei & Hamilton, 2016), the interpretation of the Cronbach's alpha coefficient for higher order models become the ratio of the total common variance explained by lower level constructs without their unique factor variance that is not accounted by the higher level factor.

hence, the underperformance of Cronbach's alpha coefficients are anticipated for STRONG-R first-order factors/sub-scales.

### Second Order Constructs/Scales

Cronbach's alpha coefficients for second order/scales performed better, ranging from .429 (*Anti-Social Influence*) to .852 (*Anti-Social Personality*). This indicates that most of the second order scales accounted for a considerable amount of common variance. This further supports the latent factor structure of the instrument while also indicates the acceptable-to-good reliability of the second order scales.

### Third Order Constructs

Third order factor/constructs performed even better and all passed the industry standard criterion of .70, indicating good scale internal consistency. The Cronbach's alpha coefficients ranged from .720 (*Anti-Social History*) to .843 (*Reintegration Needs*). Similar to the reasoning above, the relatively high Cronbach's alpha coefficients present evidence of a reliable and stable latent factor structure of the instrument.

**Table 1. Reliability Coefficients for First, Second and Third Order Constructs/Scales**

Overall Model	Omega Reliability	CFI	TLI	RMSE (95%CI)	SRMR			
<i>G-Factor Solution</i>	.998	.972	.945	.059 (.056-.062)	.021			
3 <sup>rd</sup> Order Factor {3 <sup>rd</sup> ,(2 <sup>nd</sup> ), & [1 <sup>st</sup> ] order loading	3 <sup>rd</sup> Order Cronbach's Alpha	2 <sup>nd</sup> Order Factor	2 <sup>nd</sup> Order Cronbach's Alpha	1 <sup>st</sup> Order Factor	1 <sup>st</sup> Order Cronbach's Alpha	No. of Items		
<i>Anti-Social Hx.</i> {.952}(.721) [.894]	.720	Criminal History	.677	Adult Criminal Record	.648	3		
				Prison Infraction Record	.725	4		
		Violence History	.489	DV & Violent Misd. Hx.	.445	4		
				Aggression Hx.	.184	3		
<i>Educ. &amp; Employ</i> {.968} (.848) [.714]	.723	Education & Work Experience	.695	Juvenile Record	.625	4		
				Education Level	.563	2		
		Income	.542	Work Experience	.577	2		
				Illegal Income	.408	2		
		Anti-Social Influence	.429	Legal Income	.723	Friend Anti-Social	.229	3
						Partner Anti-Social	.554	3
Family Anti-Social	.557					3		
<i>Anti-Social Propensity</i> {.984} (.720) [.877]	.770	Anti-Social Personality	.852	Deception	.582	2		
				Empathy	.639	2		
				Retreat	.636	3		
		Anti-Social Cognition	.728	Violence Propensity	.665	Egocentrism	.404	7
						Respect	.541	2
						Change	.738	2
						Thinking Error	.659	2
						Cognitive Skills	.731	2
						Physical Abusive	.582	3
						Hyper-Masculinity	.180	3
<i>Substance Abuse</i> {--}(.899)[.927]	.729	--	--	Destructive	.546	2		
				Fixation	.151	3		
				Irritability	.333	4		
				Hard Drug Use	.537	3		
Reintegration Needs {.930}(.727)[.726]	.843	Employment Barrier	.627	Drug Use & Share & Barter	.573	3		
				Drug Related Crime	.388	3		
				Physical & Mental Barrier	.237	2		
		Mental Health	.588	--	--	Necessity	.509	3
						Work Ethic	.395	3
						Systematic Barrier	.510	3
						Suicidal Propensity - History	.490	3
						Suicidal Propensity - Recent	.631	3
--	--	--	--	Reentry Needs	.843	9		

### *Highest Order Construct - Risk and Needs*

The scale of *risk-needs* consisted of five constructs, including *Anti-Social History*, *Education & Employment*, *Anti-Social Propensity*, *Substance Abuse Propensity*, and *Reintegration Needs*. These constructs are higher order factors<sup>2</sup> and congeneric<sup>3</sup>. Therefore, in the subsequent analysis, only the true score variance for each construct was used as indicators of the global risk-needs construct in the (Global) G-Factor analytic factor model<sup>4</sup>. Omega reliability was then calculated and findings are presented in Table 2 ( $\alpha = .998$ ).

**Table 2 Final Global Risk and Needs Model**

<b>Construct</b>	<b>No. of Item</b>	<b>Loading of Unweighted Total Score</b>	<b>Loading in the Final Model</b>	<b>Cronbach's Alpha</b>	<b>Model Fit Indices</b>
<i>Anti-Social History</i>	14	.445	.950	.720	<i>CFI = .972</i>
<i>Education &amp; Employment</i>	8	.549	.967	.723	<i>TLI = .945</i>
<i>Anti-Social Propensity</i>	46	.643	.984	.770	<i>df = 5</i>
<i>Substance Abuse</i>	9	.543	.899	.729	<i>RMSEA = .059</i>
<i>Reintegration Needs</i>	26	.596	.924	.843	<i>SRMR = .021</i>
<i>Global Risk-Needs</i>	103	<i>Omega Reliability = .682</i>	<i>Omega Reliability = .998</i>	<i>Cronbach's alpha = .658</i>	

### **Conclusion**

Overall, the STRONG-R scales demonstrate excellent reliability, as indicated by the omega value in the G-factor solution model (.998). The STRONG-R's lower level scales may possess as many as seven and as few as two items, which can result in substantial variation in alpha levels (Cortina, 1993). However, given the multi-dimensional aspects of the STRONG-R scales, greater weight is given to the internal consistency demonstrated by the higher order factors and G-factor solution, of which our demonstrated findings indicate substantial internal consistency and reliability.

<sup>2</sup> See Mei & Hamilton, 2016

<sup>3</sup> Loadings in Congeneric model are not statistically equivalent.

<sup>4</sup> In psychometric analysis, these variables that only consist of true score variance is referred to as phantom variables.



## REFERENCES

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Bollen, K.A. (1989). *Structural equation models with latent variables*. New York, NY: Wiley.
- Cortina, J. What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1), 98-104.
- Mei, X. & Hamilton, Z. (2016). Internal Structure of the STRONG-R. Spokane, WA: Washington State Institute for Criminal Justice.
- Thompson, M. (2004). The value of item response theory in clinical assessment: A review. *Assessment*, 18 (3), 291-307.
- Cronbach, L. (1951). Coefficient alpha and the internal structure of a test. *Psychometrika*, 16, 297-334.
- Wang, J. & Wang, X. (2012). *Structural Equation Modeling: Applications Using Mplus*. Hoboken: Wiley.
- Raykov, T. (2001). Bias of Cronbach's alpha for fixed congeneric measures with correlated errors. *Applied Psychological Measurement*, 25, 69-76.
- Joreskog, K.G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, 36, 109-133.
- Dillon, W. & Goldstein, M. (1984). *Multivariate analysis: Method and Applications*. John Wiley & Sons, Inc., New York, NY.

For further details about the STRONG-R reliability research findings, WSU Researchers can be contacted at [zachary.hamilton@wsu.edu](mailto:zachary.hamilton@wsu.edu)

Suggested Citation: Mei, X. & Hamilton, Z. (2016a). *Reliability of the STRONG-R*. Spokane, WA: Washington State Institute for Criminal Justice.