Revalidation of the Women's Risk Needs Assessment:

Institutional Results[1]

Final Report

November 2013

Patricia Van Voorhis, Ph.D.
University of Cincinnati

Ashley Bauman, M.S., M.B.A.
University of Cincinnati

Rachel Brushett, Ph.D.
University of Cincinnati

[1] This research was funded by the National Institute of Corrections under cooperative agreement 09M12GKB3. Points of view or opinions stated in this article are those of the first author and do not necessarily represent the official position or policies of the United States Department of Justice or the National Institute of Corrections.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

Revalidation of the Women's Risk Needs Assessment:
Institutional Results[2]
Final Report

November 2013

Executive Summary

Patricia Van Voorhis, Ph.D.
University of Cincinnati

Ashley Bauman, M.S., M.B.A.
University of Cincinnati

Rachel Brushett, Ph.D.
University of Cincinnati

[2] This research was funded by the National Institute of Corrections under cooperative agreement 09M12GKB3. Points of view or opinions stated in this article are those of the first author and do not necessarily represent the official position or policies of the United States Department of Justice or the National Institute of Corrections.

Introduction

In order to facilitate appropriate treatment planning and risk management for women offenders, the National Institute of Corrections and the University of Cincinnati entered into a cooperative agreement to create and validate a women's, dynamic, risk/needs assessment, the Women's Risk Needs Assessment (WRNA). Development of two types of gender-responsive assessments began in 1999 with a pilot study in the Colorado Department of Corrections and later continued with three projects in Maui, Minnesota, and Missouri. The first assessment, called the Women's Risk Needs Assessment – Trailer (WRNA-T)(or "the trailer") was designed to supplement existing dynamic risk/needs assessments such as the Level of Service Inventory-Revised (LSI-R) (Andrews & Bonta, 1995) and the Northpointe COMPAS (Brennan, Dieterich, & Oliver, 2006). The second, the Women's Risk Needs Assessment (WRNA), was an assessment that could be used on its own, as a "stand-alone," dynamic, risk/needs assessment, comprised of both gender-neutral and gender-responsive scales. Extensive literature searches and focus groups with correctional administrators, treatment practitioners, line staff, and women offenders informed both of the assessments. Both instruments contained an interview and a self-report survey. The full instrument, and many of the questions now contained on the WRNA-T, was developed by members of the Women's Issues Committee of the Missouri Department of Corrections (MDOC) in collaboration with researchers at the University of Cincinnati and key staff from the National Institute of Corrections. This construction process also benefitted from the expertise of substance abuse specialists, psychologists and other mental health professionals on staff with MDOC.

This report presents the results of a second cooperative agreement between the National Institute of Corrections (NIC) and the University of Cincinnati which was begun in 2009. The

present report focuses on the Institutional WRNA, the assessment intended for use in prison settings. In addition to revalidating the original 2008 assessment, the project sought to further refine the assessment. The reason for seeking to improve the Institutional WRNA was attributable to the fact that original tool was developed on a small sample (N=272) of Missouri inmates. In contrast, the present study secured a sample of 484 participants in two sites, Missouri and Ohio, to test gender-neutral variables and 683 participants across three sites, Missouri, Ohio, and Rhode Island to test gender-responsive variables. The following analytical steps were employed:

1. The original 2008 risk/need domains (scales measuring specific needs)(see Appendices A and C) were tested, through analysis of correlations with outcome measures. Outcome measures for the present study consisted of prison misconducts, serious prison misconducts, and aggressive prison misconducts.

2. The original stand-alone WRNA risk/needs score (total score)(see Appendix B for the scoring form), developed for the 2008 construction validation research was tested on the Missouri and Ohio research samples.

3. The cumulative 2008 WRNA-T scale (see Appendix D for the scoring form) was added to the LSI-R for the Rhode Island samples and tested for predictive validity. Tests of the WRNA-T involved computation of its incremental validity. The examination of whether the WRNA-T made a statistically significant contribution to the LSI-R focused upon the partial correlation between the total LSI-R and WRNA-T score with misconducts while controlling for the LSI-R score alone.

4. The current study collected data on a number of new test items to determine whether they improved the predictive validity of individual domain/need scales. Items were tested on a split-half sample of all prison sites studied. A construction and a revalidation sample were drawn through a systematic random selection process from a pool of all cases with at least 3 months of follow-up data. The sample of all cases with at least 3 months of follow-up was divided into a construction sample (N=244 for gender-neutral scales and 322 for gender-responsive scales) and a revalidation sample (N=240 for gender-neutral scales and 318 for gender-responsive scales). Every other case was selected for the construction sample, and the remaining cases were reserved for the revalidation sample. New scales were developed on the construction validation sample and retested (confirmed) on the revalidation sample. See Appendix F and H for the revised WRNA and WRNA-T interviews.

5. Total risk/needs scales were developed in the construction validation sample, retested on the revalidation sample, and then tested for specific sites. See Appendix G and I for the revised scoring forms. Data analysis employed bivariate correlations (Pearson's r) and analysis of receiver operating characteristics (AUC).

METHODOLOGY

A total of 792 women offenders were interviewed for inclusion in the prison study. These included 137 participants from Missouri, 399 from Ohio, and 256 from Rhode Island. However, a number of these participants were released early, prior to conclusion of the 6 and 12 month follow-up windows. Participants were not included in the follow-up data analysis unless they had served 3 months for inclusion in the analysis of 6 month data and 9 months for inclusion in the analysis of 12 month data. This reduced sample sizes for tests of predictive validity at 6 months to 98 participants from Missouri, 386 from Ohio, and 154 from Rhode Island (80.6 percent of the original sample). Of the original sample, 469 participants (59.2 percent of the original sample) were eligible for inclusion in the 12 month analysis. The size of each 12 month sample was as follows: a) Missouri (N=53); Ohio (N=347); and Rhode Island (N=69). The Missouri and Ohio samples participated in interviews and tests of the stand-alone WRNA. Rhode Island DOC utilized the WRNA Trailer (WRNA-T) with the LSI-R, and furnished data needed to validate the WRNA-T's contribution to the LSI-R.

The WRNA-T was in full use in Rhode Island at the time of this study and interviewers were case managers employed by RIDOC. Case managers also conducted interviews for the Missouri sample; however, limited cooperation from the site greatly reduced the sample size and representativeness of this sample. Finally, Ohio was a research site; all interviews were conducted by staff and advanced graduate students from the University of Cincinnati.

Sample descriptions appear in Table 1 of the main report. When the samples were partitioned into construction and revalidation samples, no significant differences were noted on social, demographic, and criminal history background variables.

Offense-Related Outcome Measures

With the exception of women who were released early, participants were followed-up at 6 and 12 month intervals on the following measures: a) any prison misconduct (yes or no); b) number of prison misconducts; c) any serious misconduct (yes or no); d) number of serious misconducts; e) any aggressive misconduct (yes or no); and f) number of aggressive misconducts. Distributions on these measures are shown by state in Table 1. Serious infractions included behaviors such as drug-related behaviors, sexual contact, various forms of contraband, gambling, refusing to accept institutional assignments, manufacturing or possessing a weapon, stealing property, tampering with firearms or locks, as well as the subset of items that were also captured by the prison aggression measures. Institutional aggression measures pertaining to women typically tapped less lethal forms of aggression such as fighting and threatening behaviors.

Table 1 also shows considerable differences in the prevalence rates of disciplinary infractions across sites. These likely reflect different staff and institutional cultures, staff inmate management skills, and the appropriateness of a facility's infraction code for women offenders. As can be seen, base rates on the aggression measure are especially limited.

Table 1. Follow-up Measures by Time Frame and Site.

Site	All Misconducts			Serious Misconducts			Aggressive Misconducts		
	N	%	$\bar{X}$	N	%	$\bar{X}$	N	%	$\bar{X}$
6 Month Follow-up									
Missouri (N=98)	67	68.4	1.29	28	28.6	0.42	2	2.0	0.04
Ohio (N=386)	64	16.6	0.24	54	14.0	0.18	30	7.8	0.10
Rhode Island-WRNA-T (N=156)	65	41.7	0.74	39	25.0	0.37	7	4.5	0.06
12 Month Follow-up									
Missouri (N=53)	40	75.5	2.87	23	43.4	0.87	3	5.7	0.11
Ohio (N=347)	93	26.8	0.52	80	23.1	0.36	51	14.7	0.20
Rhode Island-WRNA-T (N=69)	36	52.2	1.28	25	36.2	0.73	8	11.6	0.20

RESULTS

Revalidation of Original Institutional WRNA

Risk/Need Domains: Correlations between the 2008 WRNA risk/needs domains and prison misconducts are shown in Table 2, for the combined Ohio, Missouri, and Rhode Island samples. A number of the patterns found in earlier studies appear in this sample as well. First, there are several domains (e.g., educational assets and family and parenting issues) that appear to be more predictive in community settings where they more potently affect daily life than in institutional settings (see Van Voorhis et al., 2010; Van Voorhis et al., 2012; Van Voorhis et al., 2013). Another pattern finds that static criminal history items, which in most states are the central variables of prison custody classification systems, are less predictive or equally predictive to domains that described troubled women, e.g. anger, depression, recent substance abuse, and

child abuse. Finally, results for longer follow-up windows are stronger than those for short (6 month) follow-up windows. The strongest predictors for the sample as a whole consisted of criminal history, anger, recent substance abuse, and depression. It is noteworthy, however, that correlations seldom surpassed .20; however, these are considerably higher when the data are disaggregated by site (see Tables 5 through 7 in the main report).

Total Risk Scale: The original total risk scale represented the total score of all risk/need domains found to be predictive in the 2008 construction validation study. This scale summed scores for criminal history, antisocial attitudes, anger, psychosis, family conflict and collapsed measures of mental health history, depression, child abuse, substance abuse history, and relationship dysfunction. The original scale also subtracted a strength, family support (collapsed), from the total score. Ideally, correlations should surpass .27, and AUC values should surpass .70. Such results were seen for some of the Ohio tests but not for the Missouri analysis.

WRNA-T: The Trailer for the Women's Risk/Needs Assessment (WRNA-T) was comprised of gender-responsive items, i.e., anger, psychosis, and collapsed measures of mental health history, depression, family support and relationship dysfunction. For purposes of risk assessment, the trailer was not intended to be used alone but rather to be added to the LSI-R, or an alternative gender-neutral risk/needs assessment. The WRNA-T could be tested on its own in all sites. However, its contribution to an existing gender-neutral assessment could only be tested in Rhode Island, where the LSI-R was used as the gender-neutral assessment.

The gender-responsive variables by themselves offered statistically significant predictions of prison misconducts for all of the sites (see Table 4). However, results were stronger for the Rhode Island and Ohio sites than for the Missouri site. We do not apply a

Table 2. Bivariate Relationship between Original WRNA Scales and Prison Misconducts, All Sites.

	6 month (N=484 & 640)[a]						12 month (N=400 & 470)[b]					
	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N
Interview												
Criminal history	.10**		.11***	.07**	.12***	.09**	.17***	.11**	.18***	.16***	.19***	.16***
Attitudes	.06*		.09**	.06*		.06*						
Educational needs	.08*						.07*	.08*				
Educational assets[c]	-.07**	-.11***		-.07**		-.05*	-.06*	-.11***	-.07*	-.09**	-.08**	-.10**
Employment/financial[c]	.12***	.12***	.08**	.08**	.07**	.06*	.10**	.15***	.13***	.14***	.12***	.12***
Antisocial friends	.10**		.12***	.09**	.09**	.06*	.11***	.08**	.14***	.08**	.14***	.12***
Anger[c]	.16***	.14***	.17***	.17***	.28***	.26***	.24***	.19***	.25***	.24***	.32***	.32***
Mental health hist.[c]	.09***		.06*				.09**					
Depression (symptoms)[c]	.09***	.08**	.10***	.09***	.12***	.11***	.07*	.08**	.08**	.10**	.15***	.15***
Psychosis (symptoms)[c]									.07*	.10***	.09**	.07*
Child abuse (interview) [c]	.09***	.07**	.12***	.10***			.08*	.09*	.08*	.16***	.11**	.08*
Adult abuse (interview) [c]												
Sex abuse (adult or child) [c]	.06*		.08**	.07**							.06*	
Physical abuse [c]	.07**		.08**							.09*		
PTSD[c,d]	.08**	.07*	.12***	.10***	.07**	.07*	.11**	.08**	.10**	.08**	.11**	.09**
Substance abuse history							.08**		.08*			
Substance abuse (recent)[c]	.18***	.17***	.11***	.09**			.19***	.14***	.16***	.08*		
Parental involvement (strength)[c,e]	.07*	.08*			.07*							
Parental difficulties (all)		-.06*	-.06*	-.07**					-.07*	-.08**	-.07*	-.07*
Relationship satisfaction (strength) [c]												
Family conflict [c]												
Family support (strength) [c]												
Survey												
Relationship dysfunction [c]			.08*									
Self-efficacy (strength) [c]	-.06*						-.06*				-.09**	
Child abuse [c]	.07**		.13***	.09**	.05*	.05*			.07*	.08**	.11***	.10***
Adult abuse [c]	.06*											
Parental stress [c]											-.06*	

***p≤.01, **p≤.05, *p≤.10

[a]Participants were in the 6 month window for at least 3 of the 6 months. N for scales appearing on the full WRNA only=484; N for scales appearing on the trailer and the full WRNA=640.

[b]Participants were in the 12 month window for at least 9 of the 12 months. N for scales appearing on the full WRNA only=400; N for scales appearing on the trailer and the full WRNA=470.

[c] Indicates a scale appearing on both the full WRNA and the trailer (N=640 at 6 mos.)(N=470 at 12 months).

[d] Questions were not asked by all interviewers (N=535 at 6 mo.)(N=414 at 12 months).

[e] Scale pertains to mothers (N=381 at 6 mo; N=273 at 12 mo.)

Table 3. Bivariate Relationship between Original WRNA Cumulative Scales and Prison Misconducts

	6 month						12 month					
	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N
Full Original WRNA Assessment (Missouri & Ohio)(N=484 at 6 mo)(N=400 at 12 mo.)												
WRNA-Full Scale	.16***	.06	.21***	.14***	.18***	.15***	.20***	.10**	.20****	.17***	.23***	.20***
AUC	.61		.66		.79		.63		.64			
Levels	.17***		.18***	.12***	.14***	.13***	.20***	.10**	.17***	.13***	.19***	.18***
AUC	.61		.64		.66		.62		.61			
Missouri (N=98 at 6 mo., N=53 at 12 mo.)												
WRNA-Full Scale			.16*								.20**	
AUC	.55		.58		.62		.46		.41		.70	
Levels	.15*								-.23**			
AUC	.59		.57		.45		.49		.38		.56	
Ohio (N=386 at 6 mo., N=347 at 12 mo.)												
WRNA-Full Scale	.27***	.18***	.25***	.19***	.18***	.15***	.27***	.23***	.28***	.24***	.23***	.21***
AUC	.72		.72		.70		.69		.70		.69	
Levels	.25***	.19***	.22***	.19***	.16***	.16***	.26***	.23***	.25***	.22**	.21***	.21**
AUC	.69		.68		.68		.67		.67		.67	

***p≤.01
**p≤.05
*p≤.10

Table 4. Bivariate Relationship between Original WRNA Trailer Cumulative Scales and Prison Misconducts

	6 month						12 month					
	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N
Full Original WRNA Assessment (Missoui, Ohio, & Rhode Island)(N=640 at 6 mo)(N=470 at 12 mo.)												
WRNA-Trailer All Sites	.13***	.10***	.17***	.14***	.19***	.17***	.14***	.13***	.17***	.17***	..25***	.24***
Missouri (N=98 at 6 mo.)(N=53 at 12 mo.)												
WRNA-Trailer	.17**		.27***	.21***	.16*	.16*		.24**		.18*	.20**	.18*
Ohio (N=386 at 6 mo.)(N=347 at 12 mo.)												
WRNA-Trailer	.25***	.18***	.23***	.18***	.19***	.17***	.25***	.21***	.25***	.22***	.24***	.23***
Rhode Island Intake (N=154 at 6mo)(N=69 at 12 mo.)												
LSI-R Total Scale	.15**	.21***			.15**	.15**		.28***		.25**	.32***	.34***
WRNA-Trailer	.17**	.25***	.13**	.16**	.17**	.16**		.34***	.22**	.34***	.40***	.42***
LSI-R + WRNA Trailer	.17**	.24***		.10*	.18**	.17**		.32***	.17*	.30***	.37***	.39***
AUC	.59				.77				.60		.80	
Partial corr.	.12*	.17**	.14**	.14**	.11*			.25**	.18*	.26**	.29***	.32***

***p≤.01
**p≤.05
*p≤.10

standard of r=.27 and AUC of .70 to the WRNA-T, because, by itself, the WRNA-T is not intended to serve as a risk assessment.

When used as a supplement to the LSI-R in Rhode Island, the predictive merits of the WRNA-T were stronger than those for the LSI-R. As a result, the predictive validity of the combined WRNA-T and LSI-R was lower than that for the WRNA-T alone, but considerably higher than correlations for the LSI-R, alone. Moreover, on most outcome measures, the variation attributable to the WRNA-T alone (partial correlation) was significant.

Revision of the Total Risk Scale

Test questions and item analysis resulted in changes to 8 scales, criminal history, antisocial friends, depression, recent substance abuse, relationship stability, relationship dysfunction, family conflict, and family support. Tests of the revised scales are shown in Table 5, below for the construction and revalidation scales. Results for specific sites are shown in Appendix J.

Unfortunately, some scales remain in need of improvements. For example, the structure and context of the family support and family conflict scales need to be improved through additional research. In addition, it is likely that the relationship scales will continue to be sample specific.

Revisions to the above scales also improved the predictive validity of the total risk/needs scale over results shown for the 2008 WRNA. The risk scale for the standalone WRNA and the WRNA-T summed the risk/need factors found to be significantly related to institutional misconducts, and then subtracted strengths. The factors included in each scale are shown below:

Table 5. Bivariate Relationship between Revised WRNA Scales and Prison Misconducts, Construction and Revalidation Samples.

	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N
Construction Sample												
	6 month (N=244/322)[a]						**12 month (N=209/240)[b]**					
Criminal history	.25***	.23***	.25***	.21***	.18***	.18***	.31***	.30***	.35***	.33***	.25***	.25***
Antisocial friends	.11**		.14***	.11**	.10*	.10*	.11*		.13**		.16***	.13**
Depression (symptoms) Recoded	.11**	.09**	.11**	.09*	.14***	.14***			.09*	.10**	.18***	.19***
Substance abuse (recent)	.22***	.22***	.14***	.13**			.22***	.20***	.15***	.12**		
Relationship stability (strength)	-.13***	-.10**	-.09**	-.11**	-.13***	-12**	-.08*		-10*	-.10*	-.17***	-.16***
Family conflict[c]	.12**		.09*									
Family support (recoded)	.09*											
Relationship dysfunction	.11**	.07*	.16***	.10**							.18***	.10*
Revalidation Sample												
	6 month (N=240/318[a]						**12 month (N=191/230)[b]**					
Criminal history							.17***	.14**	.24***	.21***	.17***	.11*
Antisocial friends	.09*		.11**	.09*			.13**	.13**	.17***	.12**	.12**	
Depression (symptoms) Recoded		.08*	.14***	.13***	.14***	.13***		.10*	.13**	.12**	.15***	.14**
Substance abuse (recent)	.18***	.13***	.09*				.18***	.14**	.18***			
Relationship stability (strength)	-.07*				-.08*					-.09*	-.11**	-.10*
Family conflict[d]			.16***	.13**			.12*	.15**	.12*	.20***		.10*
Family support (strength)												
Relationship dysfunction												

***p≤.01, **p≤.05, *p≤.10

[c]Participants were in the 6 month window for at least 3 of the 6 months.

[b]Participants were in the 12 month window for at least 9 of the 12 months.

[c] Rhode Island cases are omitted, because interviewers could not ask test questions. (N=244 at 6 mo; N=209 at 12 mo)

[d] Rhode Island cases are omitted, because many interviewers did not ask the test questions. (N=240 at 6 mo; N=191 at 12 mo)

WRNA Stand Alone
Criminal history
Anger
Antisocial friends
Recent substance abuse
Depressionsymptoms (collapsed)
Psychotic symptoms
Child abuse
Relationship dysfunction
Family support

WRNA Trailer
Anger
Depression symptoms (collapsed)
Psychotic symptoms
Child abuse
Relationship dysfunction
Family support

Family support is subtracted from this total and appears in Part II of the scoring form.

Predictive merits of the revised standalone scale are seen in Table 6 for the Missouri and Ohio sites where it was tested. Correlations and AUC values were especially strong for the combined sample at the 12 month point, and for predictions of serious and aggressive misconducts. Additionally, AUC values for the Ohio sample were at or above .70 on all outcome measures, regardless of time frame. The Missouri results showed strong predictions of serious misconducts at 6 months, but results became much less stable at the 12 month point when the sample was considerably reduced in size.

Table 6 also shows that the gender responsive risk scale was more predictive, in most tests, than other models available to correctional agencies, including static custody scales and gender-neutral risk/needs scales.

This comparison is also seen in Table 7, where the results of augmenting the LSI-R with the WRNA-T are shown for the Rhode Island sample. The WRNA-T alone is significantly correlated with prison misconducts in 11 of 12 tests and makes a statistically significant contribution (partial correlation) to the LSI-R in 9 of 12 tests. Finally, the WRNA-T alone is significantly correlated with prison misconducts for the sample as a whole and the Ohio sample in all tests and in 8 of 12 tests conducted on the Missouri sample.

Table 6. Bivariate Relationship between Revised Cumulative Scales and Prison Misconducts, Ohio and Missouri.

	6 month						12 month					
	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N
Full Revised WRNA Assessment, (Missouri & Ohio (N=484/400)												
Static Custody	.16***	.14***	.16***	.15***	.11***	.09**	.25***	.23***	.29***	.28***	.21***	.19***
AUC	.61		.61		.62		.65		.69		.67	
Gender Neutral	.21***	.16***	.20***	.17***	.11***	.08**	.26***	.22***	.30***	.24***	.21***	.18***
AUC	.63		.65		.62		.65		.60		.68	
WRNA Stand Alone	.21***	.16***	.26***	.21***	.18***	.14***	.26***	.21***	.30***	.26***	.27***	.23***
AUC	.64		.70		.71		.67		.70		.73	
WRNA Stand Alone-Levels	.22***	.17***	.26***	.21***	.19***	.15***	.29***	.22***	.31***	.27***	.29***	.24***
AUC	.63		.68		.69		.66		.68		.71	
Missouri (N=98/53)												
Static Custody	--	--	--	--	--	--	--	.18*	.27**	.20*	--	--
AUC	.51		.51				.57		.64			
Gender Neutral	.21**	.13*	.21**	.15**	--	--	.20*	.18*	.25**	--	--	--
AUC	.62		.63				.63		.65			
WRNA Stand Alone	.25***	.18**	.33***	.26***	--	--	.18*	.22*	.18*	.19*	--	--
AUC	.65		.70				.62		.59			
WRNA Stand Alone-Levels	.20**	.18**	.28***	.22***	--	--	--	.20*	--	--	--	--
AUC	.62		.67		--		--		--		--	
Ohio (N=386/347)												
Static Custody	.21***	.18***	.19***	.18***	.13***	.11***	.26***	.27***	.29***	.29***	.24***	.23***
AUC	.65		.65		.63		.66		.69		.68	
Gender Neutral	.21***	.17***	.20***	.17***	.13***	.11***	.26***	.26***	.30***	.27***	.25***	.23***
AUC	.66		.61		.65		.66		.70		.70	
WRNA-Stand Alone	.25***	.20***	.24***	.21***	.19***	.16***	.30***	.27***	.33***	.29***	.30***	.26***
AUC	.70		.71		.71		.70		.73		.74	
WRNA Stand Alone-Levels	.26***	.19***	.25***	.21***	.22***	.18***	.32***	.27***	.34***	.30***	.32***	.28***
AUC	.68		.69		.71		.69		.71		.73	

***p≤.01
**p≤.05
*p≤.10

Table 7. Bivariate Relationship between Revised Cumulative WRNA Trailer Scales and Prison Misconducts.[a]

	6 month						12 month					
	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N
WRNA Trailer, (Missouri, Ohio & Rhode Island (N=640/400)												
Cumulative WRNA Trailer[a]	.13***	.11***	.18***	.15***	.19***	.17***	.12***	.11***	.17***	.17***	.26***	.22***
Missouri (N=98/53)												
Cumulative WRNA Trailer[a]	.21**	.17**	.36***	.30***	.15*	.15*	--	.20*	--	.20*	--	--
Ohio (N=386/347)												
Cumulative WRNA Trailer[a]	.23***	.18***	.23***	.19***	.19***	.16***	.23***	.19***	.25***	.21***	.24***	.21***
Rhode Island (N=154/69)												
LSI-R[b]	.15**	.21***	--	--	.15**	.15**	--	.28***	--	.25**	.32***	.34***
AUC	.54		.54		.74						.79	
Cumulative WRNA Trailer[a]	.19**	.26**	.11*	.14**	.16**	.16**	--	.31***	.22**	.32**	.44***	.47***
LSI-R + WRNA Trailer[c]	.17**	.25***			.17**	.17**	--	.31***	.17*	.29***	.38***	.40***
AUC	.54		.55		.77				.59		.80	
Partial WRNA Trailer[d]	.13*	.18**	.12*	.12*			--	.19*	.17*	.22**	.33***	.36***

***p≤.01
**p≤.05
*p≤.10

[a] The WRNA Trailer is not intended to be a standalone risk assessment instrument, but rather a supplement to a gender neutral dynamic risk assessment, such as the LSI-R. Therefore, these correlations are not expected to be as high as a full risk/needs assessment.

[b] The LSI-R is a standalone risk assessment instrument correlations with outcomes at 12 months should be above .25.

[c] The LSI-R plus the WRNA trailer is a full risk/needs assessment. Correlations at 12 months should be above. 25. In some instances these are pulled down by low LSI-R values.

[d] This partial correlation shows the predictive validity (variation) that is attributable to the WRNA. It can be seen than in most instances, the WRNA makes a statistically significant contribution to that offered by the LSI-R, alone.

CONCLUSION

The applications of the revised Institutional WRNAs differ somewhat from the Probation and Prerelease WRNAs. A number of facilities currently using the Institutional WRNA do not use its total risk scale for purposes of assigning inmates to levels of custody, because the scale is comprised of a number of needs pertinent to mental health and abuse. Understandably, correctional official are reluctant to raise custody levels according to criteria over which inmates have little control. As a result, most prisons use the WRNA as a needs assessment and benefit from identifying and programing for needs that place women at risk of making poor institutional adjustments. Thus, even without formulating an overall prison risk scale, use of the tool as a needs assessment can serve dual goals of inmate behavioral change and prison safety.

Even so, revalidation tests of the 2008 WRNA and WRNA-T, with no changes, produced acceptable results for 2 of 3 research sites, Ohio and Rhode Island. Results for the Missouri prison site were unacceptable. As noted, in the methodology section, however, it was difficult to secure the cooperation of prison interviewers and potential participants.

The present study tested a number of improvements to dynamic risk/need scales which resulted in improvements to the predictive validity of these specific risk/need scales as well as to the predictive validity of the assessment's institutional risk scale. Improvements were seen even for the Missouri sample; however, they were not as conclusive as those seen for Rhode Island or Ohio.

The study has succeeded in producing a somewhat shorter assessment than the original tool. This occurs primarily with the omission of two abuse survey scales measuring adult abuse (victimization) and child abuse. Most importantly, we have much more confidence in the stability of the assessment, because it now is seen to be predictive across several jurisdictions.

Because, revisions were made using construction and revalidation samples, we have reduced the urgency for revalidation research. Nevertheless, revalidation tests by other researchers are strongly encouraged.

This study also afforded an opportunity to prepare a trailer (WRNA-T) for use with the LSI-R. In most tests, this tool significantly augmented the predictive validity of the LSI-R and provided a means for screening according to gender-responsive needs that are not contained on the LSI-R. A number of jurisdictions have chosen to use the WRNA-T solely as a needs assessment, thus avoiding the complication of adding the gender-responsive scales to the LSI-R and recalibrating risk levels. While that is a reasonable possibility, it was clear that the contribution of the WRNA-T to the predictive validity of the LSI-R was favorable (see Table 7).

Notwithstanding these contributions, there are precautions to be taken in understanding these findings. Follow-up time frames are limited to 12 months, both by the terms of NIC funding and by the releases of inmate participants. It is well known that longer follow-up time frames tend to produce better results, especially when they extend to a 24 month window. In addition, results varied somewhat from interviewer to interviewer. Separate analyses found that some interviewers produced data which achieved lower predictive validity coefficients, and missing data than others, especially on sensitive scales pertaining to abuse, trauma, and relationships. Finally, we note that interviewers for the WRNA assessments were trained immediately prior to data collection. In contrast, the Rhode Island LSI-R interviewers had been trained several years prior to this study. State officials observed that many of the LSI-R interviewers were due to receive refresher training. The difference in proximity to training may explain why the WRNA-T was somewhat more predictive than the LSI-R. Dynamic assessments

such as the WRNA and the LSI-R require careful monitoring for quality assurance; the validity of either assessment is likely to diminish when quality assurance becomes lax.

This report is the last of 6 reports prepared for NIC. Separate reports were prepared for Iowa, Rhode Island, and Missouri as their data became available. In addition, three reports summarized findings for probation, prerelease, and prison settings. Each summary report produced a final instrument considered to be optimal for that setting. A total of 11 sites were involved. Only two of these, the Ohio probation and the Missouri prison samples, produced questionable findings. Both involved limited cooperation from staff in the respective agencies. The assessments were valid in all other sites as well as in the 3 sites studied during the 2008 construction validation research.

Revalidation of the Women's Risk Needs Assessment:

Institutional Results[3]

Final Report

November 2013

Patricia Van Voorhis, Ph.D.
University of Cincinnati

Ashley Bauman, M.S., M.B.A.
University of Cincinnati

Rachel Brushett, Ph.D.
University of Cincinnati

[3] This research was funded by the National Institute of Corrections under cooperative agreement 09M12GKB3. Points of view or opinions stated in this article are those of the author and do not necessarily represent the official position or policies of the United States Department of Justice or the National Institute of Corrections.

Acknowledgements

The present report is not solely the work of the three authors, and we would be remiss if we did not thank the many individuals who helped in its creation. Behind the scenes were funders, leaders of correctional agencies, interviewers, trainers, UC research staff, and research and planning staff at each of the research sites. All worked hard to complete this research project.

First, as noted throughout the report, this study was funded through a cooperative agreement between the University of Cincinnati and the National Institute of Corrections. We are especially appreciative of Maureen Buell's patience and constructive support throughout this process.

Correctional leaders at all of the sites provided access, vision and a good deal of support to our efforts. These included A.T Wall, Director of the Rhode Island Department of Corrections (RIDOC); Roberta Richman, Assistant Director (RIDOC): Carole Dwyer, Warden (RIDOC); Al Leach, Deputy Warden, (RIDOC); Terry Foley, Professional Services Coordinator (RIDOC); Ken Findley, Professional Services Coordinator (RIDOC); David Rost, Deputy Director of the Missouri Department of Corrections (MDOC); Tom Hodges, Chief State Supervisor of Probation and Parole (MDOC); Julie Kempker, Assistant Division Director of Probation and Parole (MDOC); Matt Sturm, Director of Offender Rehabilitation Services (MDOC); Marta Nolin, Assistant Division Director for Substance Abuse Services (MDOC); Jim Wiseman, Chief of Staff Training (MDOC); Tina Waldron, Former ReEntry Manager (MDOC); Cyndi Prudden, Deputy Director of Adult Institutions (MDOC); Ginine Trim, Warden (ODRC); Elizabeth Wright, Administrative Assistant (ODRC); and Steve Van Dine, Chief of Bureau of Research and Evaluation (ODRC). In many cases, these leaders participated in planning sessions that contributed tremendous wisdom to our efforts.

Planning teams included the Women's Issues Committee of the Missouri Department of Corrections. Together with staff from the National Institute of Corrections and the University of Cincinnati, they authored the scales in 2003. Their contribution to this project was essential and the results speak highly of their expertise. A more contemporary team in Missouri, the Gender Responsive Assessment Implementation Team (GRAIT) was responsible for guiding the WRNA's implementation statewide.

The project depended upon the assistance of staff in research and planning departments. We worked extensively with Bree Derrick and Erin Boyar at RIDOC, David Oldfield and Fred Martin at MDOC, and Brian Kowalski and John Chin at ODRC. These individuals electronically compiled assessment and follow-up data for the project. We are grateful to them for the time, patience, and commitment they offered to this project.

This project could not have been completed without the dedication of interviewers and their trainers. My staff at the University of Cincinnati, including Ashley Bauman, Krista Gehring, Valerie Bell and Rachel Brushett have logged many travel hours, prepared the training curriculum, and have effectively trained assessors in over 20 sites. The project could not have been completed without the dedication of correctional personnel who conducted interviews and

in some cases provided case management to women offenders benefitting from the interviews. Their work serves as a model to others in the field. I can certainly thank the UC interviewers in this regard: Drs. Valerie Bell, Krista Gehring and William Stadler, as well as Ashley Bauman, Brittany Groot, Beth Ellefson, Noreen Loftus-Spilman, and Lindsay Morrow.

We are especially appreciative to the nearly 800 women offenders who participated in this study. Their answers to assessment questions are the foundation of this work. Without the candid and honest disclosure of their life stories, this project would not have been possible. Moreover, many lent valuable input into the design of the assessment. They contributed their time willingly and without any remuneration. The future beneficiaries of this work owe them their gratitude.

Patricia Van Voorhis, Ph.D.
Ashley Bauman, M.S., M.B.A.
Rachel Brushette, Ph.D.

I am indebted to Ashley Bauman, Project Manager, for her effective work in disseminating the WRNA and her skill in managing project activities. Ashley began her work on the WRNA projects as a graduate student in the School of Criminal Justice. She quickly became the project's lead trainer and change agent. That the WRNA has been successfully adopted by a rapidly growing number of jurisdictions is primarily attributable to Ashley. She is a trusted presence in many state and local correctional agencies and has guided large-scale implementations with patience, talent and grace.

Pat Van Voorhis

Introduction

By the late 1990s, a number of scholars voiced concern for the applicability of the current generation of risk/needs assessments to women offenders. By then, dynamic risk/needs assessments had been widely adopted to address both security and treatment needs of correctional clientele. These assessment tools served the function of classifying offenders according to low, medium and high risk to assist agencies in managing the security needs of offenders. Additionally, they identified the needs or risk factors that were likely to contribute to offender recidivism. In doing so, these assessments also identified programmatic needs of offenders. Unfortunately, most of the widely used risk and need assessments were created for men and later applied to women prior to an examination of their appropriateness or validity (Bloom, Owen, & Covington, 2003; Chesney-Lind, 1997; Morash, Bynum, & Koons, 1998; Van Voorhis & Presser, 2001). Most importantly the assessments ignored needs central to women including: relationships, mental health problems, parental and childcare issues, safety, poverty, abuse and victimization, and strengths pertaining to family support, relationship support, self-efficacy, and educational attainments (Blanchette, 2004; Blanchette & Brown, 2006; Brennan, 1998; Brennan & Austin, 1997; Farr, 2000; Reisig, Holtfreter, & Morash, 2006; Van Voorhis, Wright, Salisbury & Bauman, 2010; and Van Voorhis, 2012).

To remedy this situation and other problems created by the lack of gender-responsive assessments, the National Institute of Corrections and the University of Cincinnati entered into a cooperative agreement to create and validate a women's, dynamic, risk/needs assessment, the Women's Risk Needs Assessment (WRNA). Development of two types of gender-responsive assessments began in 1999 with a pilot study in the Colorado Department of Corrections and later continued with three projects in Maui, Minnesota, and Missouri. The first assessment,

called the Women's Risk Needs Assessment – Trailer (WRNA-T)(or "the trailer") was designed to supplement existing dynamic risk/needs assessments such as the Level of Service Inventory-Revised (LSI-R) (Andrews & Bonta, 1995) and the Northpointe COMPAS (Brennan, Dieterich, & Oliver, 2006). The second, the Women's Risk Needs Assessment (WRNA), was an assessment that could be used on its own, as a "stand-alone," dynamic, risk/needs assessment, comprised of both gender-neutral and gender-responsive scales. Extensive literature searches and focus groups with correctional administrators, treatment practitioners, line staff, and women offenders informed both of the assessments. Both instruments contained an interview and a self-report survey. The full instrument, and many of the questions now contained on the WRNA-T, was developed by members of the Women's Issues Committee of the Missouri Department of Corrections (MDOC) in collaboration with researchers at the University of Cincinnati and key staff from the National Institute of Corrections. This construction process also benefitted from the expertise of substance abuse specialists, psychologists and other mental health professionals on staff with MDOC.

The Women's Risk Needs Assessment (WRNA) was informed by two perspectives on offender rehabilitation: a) research by Canadian scholars Donald Andrews, Paul Gendreau, James Bonta, and others, which stressed the importance of assessing and treating dynamic risk factors (see Andrews & Bonta, 2010; Gendreau, Little & Goggin, 1996); and b) scholarship by feminist criminologists (e.g., Belknap, 2007; Bloom et al., 2003; Chesney-Lind, 1997; Daly, 1992; Morash, 2006; 2010) stressing the importance of women's unique "pathways" to crime. Both perspectives were relevant to the importance of programming targeted to dynamic risk factors for women offenders. However, the pathways perspective asserted that women's unique

needs were not adequately tapped by the current generation of risk/needs assessments, such as the LSI-R and the COMPAS.

The construction validation studies also produced different versions for specific types of correctional populations, because it was discovered that the predictive validity of both the gender neutral and the gender-responsive variables varied by correctional settings, e.g., prerelease, probation, and prisons. The rationale for three different assessments reflects evidence that risk factors for women differ across correctional settings (Van Voorhis, Wright, Salisbury, & Bauman, 2010). The needs that are predictive of prison misconducts are not always the same as those that predict new offenses committed by probationers and parolees. These differences required that different risk/need scales be summed in the course of calibrating the total risk scores that produced the classification of high, medium, and low risk. However, the assessment itself was the same for each setting.

In 2009, the National Institute of Corrections (NIC) entered into a second cooperative agreement with the University of Cincinnati which produced the present study. Since the earlier assessments were created through construction validation, a key goal of the present study was to revalidate the original versions on new samples of offenders to assess the level of shrinkage in predictive validity from the construction to revalidation studies.

The 2009 cooperative agreement also sought to refine several of the dynamic risk/needs scales in order to further improve predictive validity. This round of research tested a number of new items, listed on the assessment as "case management questions", that allowed for the exploration of their potential contributions to a revised assessment. Of course, creating a revised assessment also required another revalidation. To accomplish that, the new studies furnished

larger samples than produced by the 2008 research and afforded an opportunity to partition the combined samples into construction and revalidation samples.

The present report focuses on the Institutional WRNA, the assessment intended for use in prison settings. In addition to revalidating the original 2008 assessment, the project sought to further refine the assessment. The reason for seeking to improve the Institutional WRNA was attributable to the fact that original tool was developed on a small sample (N=272) of Missouri inmates. In contrast, the present study secured a sample of 484 participants in two sites, Missouri and Ohio, to test gender-neutral variables and 683 participants across three sites, Missouri, Ohio, and Rhode Island to test gender-responsive variables. The specific goals of these scale revisions are as follows:

1. To test the contributions of new items to the predictive validity of specific risk/needs scales as well as to the total risk scale representing the cumulation of risk/need factors predictive of prison misconducts.

2. To assure that those scales were valid on samples that were not part of the construction of the new scales. In other words to revalidate the revised scales through a split half validity test.

3. To produce an assessment that was more likely to work across samples and not be sample specific. Up to this point, it has been necessary to validate the WRNA on specific samples as data became available (see Van Voorhis et. al, 2010). While it has been advantageous to jurisdictions have an assessment specifically tailored to their use, the process resulted in slightly different total scales for each sample. The intent of the present study was to develop a single, more universal, assessment that would be applicable to prisons.

4. To develop a trailer for the LSI-R. The 2004-2008 construction validation study did not finalize a supplement to the LSI-R. It was possible to do so in the present study.

The applications of the revised Institutional WRNAs differ somewhat from the Probation and Prerelease WRNAs. A number of facilities currently using the Institutional WRNA do not use its total risk scale for purposes of assigning inmates to levels of custody. As will be seen, the institutional risk scale is comprised of a number of needs pertinent to mental health and abuse,

because the study finds them predictive of serious prison misconducts. Understandably, correctional official are reluctant to raise custody levels on criteria over which inmates have little control. As a result, most prisons use the WRNA as a needs assessment and benefit from identifying and programing needs that place women at risk of making poor institutional adjustments. Even without formulating an overall prison risk scale, use of the tool as a needs assessment can serve dual goals of inmate behavioral change and prison safety.

Description of Participating Jurisdictions

Missouri

Under the previous cooperative agreement, the Missouri Department of Corrections (MDOC) collaborated with UC and NIC in the creation of the WRNA. In the spring of 2010 the state began a pilot project to implement the assessment. Staff participated in a 3 day training conducted by UC trainers. Select institutional caseworkers at Women's Eastern Reception, Diagnostic, and Correctional Center (WERDCC – also known as Vandalia) administered the stand alone, institutional WRNA to an intake sample of female offenders 30-45 days after their entry into the facility. While assessments were to be conducted on all women entering WERDCC over a 5 month period, it became clear that only some staff conducted the assessments for the entire time period. Other staff members appear to have conducted the interviews for a brief period of time and then discontinued its use. Unfortunately this discrepancy in protocol was not identified until study completion. Thus, it is not clear that the Missouri sample is representative of an intake cohort of Missouri female prisoners. Additionally, the failure to assess all women entering WERDCC resulted in a much lower sample size than anticipated. A total of 137 cases

(instead of the projected 400 cases) were collected during the research time period and the response rate could not be determined.[4]

Ohio

Data for this research-only sample were collected at the Ohio Reformatory for Women (ORW) over the course of one month utilizing the standalone version of the WRNA. The interview team consisted of six graduate students and one UC staff person. One of the graduate student interviewers also served as a trainer on the WRNA, and she conducted a 3 day training for all interviewers. The interview team was provided with a randomized list of women who had been at the institution at least one month and had at least one year remaining on their sentence. The women were asked to report to the administration building where researchers proceeded to ask each client in a private room if she would be interested in participating in the study. Interested participants would sign informed consent forms and then complete the WRNA. Many of the women who refused to participate in the study reported one of two reasons for refusal: 1) they were concerned they would miss other prison activities that were important to them (i.e., commissary, recreation time, etc.), or 2) they had been inundated by a large number of research projects by area universities in recent months and were tired of being involved in research studies. In total, 400 women agreed to participate in the study, and there were 106 refusals for a total response rate of 79% for this sample. [5]

[4] The Missouri study was reviewed and approved (#10122703) by the Institutional Review Board (IRB) at the University of Cincinnati in February 2011. Re-approval was granted in February 2012 and 2013.

[5] IRB approval was granted for this study (#09102901) in December 2009. Re-approvals have been granted annually.

Rhode Island

The Rhode Island Department of Corrections houses female clients at all custody levels and also acts as county jail for most of Rhode Island. On average, women in this institution served sentences of 9 months. The RIDOC began conducting the WRNA-T assessment as a supplement to the LSI-R in spring of 2008 for all women in the facility who had received prison sentences of 6 months or longer. RIDOC utilized the results for case and release planning purposes. Case managers asked women within the facility who had an LSI-R and a WRNA-T in their files if they would consent to release their assessment and file data to the University for the purposes of this study. The case managers reported a 100% response rate. A total of 223 assessments were provided by RIDOC for the study.[6]

Two of these sites, Missouri and Rhode Island, received site specific reports prior to the preparation of the present study (see, Van Voorhis, Brushett, & Bauman, 2012; Van Voorhis, Bauman, & Brushett, 2012). Ohio participated as research site only and therefore is receiving only the present report.

METHODOLOGY

A total of 792 women offenders were interviewed for inclusion in the prison study. These included 137 participants from Missouri, 399 from Ohio, and 256 from Rhode Island. Some of the participants were released early, prior to conclusion of the 6 and 12 month follow-up windows. To adjust for this, we did not include participants in the follow-up data analysis unless they had served 3 months for inclusion in the analysis of 6 month data and 9 months for inclusion in the analysis of 12 month data. With this adjustment, the analysis of predictive

[6] IRB approval for the Rhode Island study (#09120704) was granted in January 2010. Re-approval was granted in January 2011, 2012, and 2013.

validity at 6 months included 98 participants from Missouri, 386 from Ohio, and 154 from Rhode Island. Together these participants represented 80.6 percent of the original sample. Of the original sample, 469 participants (59.2 percent of the original sample) were eligible for inclusion in the 12 month analysis. The size of each 12 month sample was as follows: a) Missouri (N=53); Ohio (N=347); and Rhode Island (N=69). The Missouri and Ohio samples participated in interviews and tests of the stand-alone WRNA. Rhode Island DOC utilized the WRNA Trailer (WRNA-T) with the LSI-R, and furnished data needed to validate the WRNA-T's contribution to the LSI-R.

Sample Descriptions

Table 1 presents demographic and criminal history characteristics for the two samples that examined the WRNA stand-alone assessment (Missouri and Ohio) and the one sample that tested the WRNA-T (Rhode Island).

As can be seen, in Table 1, there were notable distinctions between the samples, especially with respect to current offenses and offense histories. The Ohio sample evidenced a much higher proportion of serious offenders, offenders incarcerated for a violent offense (44.8 percent vs. 15.3 percent for Missouri and 26.6 percent for Rhode Island). This may reflect the fact that the Ohio and Rhode Island samples were more representative of standing prison populations. As noted above, Missouri officials experienced difficulties securing the cooperation of institutional personnel, and it was not clear how their sample was drawn. Even so, Missouri population characteristics shown in Table 1 are similar to those seen in the 2008 construction validation sample (Van Voorhis et. al., 2010).

In all samples, drug offenses comprised the modal offense category, however, the proportion of convicted drug offenders was slightly higher for Missouri (30.6 percent) and Rhode Island (26.6 percent), than for Ohio (22.8 percent). Additionally, prior histories were more extensive for the Missouri sample than for Ohio. Data on prior histories were not available for the Rhode Island sample.[7]

With the exception of age (the average age was similar for each site), the samples also showed slight differences across social and demographic characteristics. The majority of participants in each sample were white; however, the Ohio sample evidenced proportionately more African American participants (25.5 percent) than the Missouri and Rhode Island samples, where African American participants comprised 19.4 percent, and 18.3 percent, respectively.

Ohio participants were more likely to be married at the time of their admission (26.9 percent), than participants in the other two sites. They were also more likely to be employed prior to their offense (74.6 percent) and to have a high school education or a G.E.D.

Table 1. Frequency and Percent Distribution of Demographic and Offense-Related Background Characteristics of Samples Participating in the Validation Study.[1]

	Missouri		Ohio		Rhode Island	
Characteristic	**N**	**Percent**	**N**	**Percent**	**N**	**Percent**
	98	100.0	386	100.0	154	100.0
Age						
18-20 years old	3	3.1	22	5.7	7	4.5
21-30 years old	30	30.6	148	38.5	50	32.5
31-40 years old	34	34.7	105	27.2	55	35.7
41-50 years old	26	26.5	83	21.5	31	20.1
51 years and older	5	5.1	28	7.3	11	7.1
	$\overline{X}$ = 35.8 yrs.		$\overline{X}$ = 34.7 yrs.		$\overline{X}$ = 34.9 yrs.	

Table Continues

[7] UC researchers had planned to secure criminal history data from the LSI-R, however, data on individual items were not available. We secured only the LSI-R domain scores from RIDOC.

Table 1. Frequency and Percent Distribution of Demographic and Offense-Related Background Characteristics of Samples Participating in the Prison Validation Study, continued.[1]

	Missouri		Ohio		Rhode Island	
Characteristic	**N**	**Percent**	**N**	**Percent**	**N**	**Percent**
	98	100.0	386	100.0	154	100.0
Race			N=385		N=131	
Asian	3	3.1	0	0.0	0	0.0
African American	19	19.4	98	25.5	24	18.3
Hispanic/Latina	1	1.0	10	2.6	13	9.9
Native American	1	1.0	5	1.3	2	1.5
White	74	75.5	258	67.0	91	69.5
Other	0	0.0	14	3.6	1	.8
Currently Married					N=149	
Yes	18	18.4	104	26.9	27	18.1
Client Have Children Under 18	N=97		N=377		N=149	
Yes	68	70.1	231	61.3	81	54.4
Employment					N=149	
Employed (full or part-time, child care, student, or disabled)	64	65.3	288	74.6	72	48.3
Not employed	34	34.7	98	25.4	77	51.7
H.S. Grad or GED			N=385		N=149	
Yes	59	60.2	259	66.7	62	41.6
Most Serious Present Offense						
Arson	1	1.0	9	2.3	1	0.6
Assault	4	4.1	28	7.3	14	9.1
Burglary	8	8.2	40	10.4	12	7.8
Damage property	7	7.1	0	0.0	0	0.0
Dangerous drugs	30	30.6	88	22.8	41	26.6
DWI	0	0.0	11	2.8	6	3.9
Family offenses	2	2.0	14	3.6	1	0.6
Forgery/Fraud	14	14.3	16	4.2	6	3.9
Homicide	3	3.1	39	10.1	4	2.6
Manslaughter	1	1.0	17	4.4	6	3.9
Kidnapping	1	1.0	1	0.2	0	0.0
Larceny	19	19.4	45	11.7	21	13.6
Loitering/soliciting (prostitution)	0	0.0	2	0.5	4	2.6
Robbery	1	1.0	54	14.0	15	9.7
Sexual assault	2	2.0	11	2.8	0	0.0
Stolen property	2	2.0	0	0.0	0	0.0
Traffic offenses	1	1.0	0	0.0	2	1.3
Weapon offenses	0	0.0	0	0.0	0	0.0
Other	1	1.0	11	2.8	21	13.6

Table Continues.

Table 1. Frequency and Percent Distribution of Demographic and Offense-Related Background Characteristics of Samples Participating in the Prison Validation Study, continued.[1]

	Missouri		Ohio		Rhode Island	
Characteristic	**N**	**Percent**	**N**	**Percent**	**N**	**Percent**
	98	100.0	386	100.0	154	100.0
Present Offense Violent						
Yes	15	15.3	173	44.8	41	26.6
Prior Felonies						
Yes	68	66.4	155	40.2	--	--
Prior Incarcerations						
Yes	53	54.1	143	37.0	--	--

[1] Inmates with at least 3 months of follow-up.

Data Analysis Plan

The goals of the present study involved validating the original 2008 Prison WRNA (see Appendices A-D) and then examining ways to improve separate risk/need domains and the total risk/needs scale. The final dynamic risk/needs scale, to be used for overall risk assessment, was the sum of individual risk/needs scales determined to be associated with prison misconducts. Two such scales were examined, a stand-alone WRNA and a WRNA Trailer (WRNA-T). The WRNA-T was designed as a gender-responsive supplement to gender-neutral risk assessments, such as the Northpointe COMPAS or the LSI-R. The present study tested a trailer for the LSI-R.[8]

[8] In order to avoid use of redundant scales, the composition of the WRNA T was specific to the gender-neutral assessment being used. For example, the Northpointe COMPAS did not contain mental health scales. Therefore, the COMPAS WRNA-T includes all of the WRNA Mental Health Scales, Mental Health History, Depression/Anxiety, and Psychosis. In contrast, the LSI-R, has a global mental health scale---Emotional/Personal. Therefore the LSI-R Trailer included only the Depression/Anxiety Scale and the Psychosis Scale of the WRNA and not the Mental Health History scale.

The following analytical steps were employed:

1. The original, risk/need domains (scales measuring specific needs)(see Appendices A and C) were tested, through analysis of correlations with outcome measures. Outcome measures for the present study consisted of prison misconducts, serious prison misconducts, and aggressive prison misconducts. These tests involved the same items and scoring protocols resulting from the 2008 construction validation study.

2. The original stand-alone WRNA risk/needs score (total score)(see Appendix B for the scoring form), developed for the 2008 construction validation research was tested on the Missouri and Ohio research samples.

3. The cumulative 2008 WRNA-T scale (see Appendix D for the scoring form) was added to the LSI-R for the Rhode Island samples and tested for predictive validity. Tests of the WRNA-T involved computation of its incremental validity. The examination of whether the WRNA-T made a statistically significant contribution to the LSI-R focused upon the partial correlation between the total LSI-R and WRNA-T score with misconducts while controlling for the LSI-R score alone.

4. The current study collected data on a number of new test items to determine whether they improved the predictive validity of individual domain/need scales. Items were tested on a split-half sample of all prison sites studied. A construction and a revalidation sample were drawn through a systematic random selection process from a pool of all cases with at least 3 months of follow-up data. The sample of all cases with at least 3 months of follow-up was divided into a construction sample (N=244 for gender-neutral scales and 322 for gender-responsive scales) and a revalidation sample (N=240 for gender-neutral scales and 318 for gender-responsive scales). Every other case was selected for the construction sample, and the remaining cases were reserved for the revalidation sample. New scales were developed on the construction validation sample and retested (confirmed) on the revalidation sample. A description of each of the samples is shown on Table 2. Table 2 shows very similar distributions across samples with no significant differences between the samples on any of the social and criminal history background characteristics. There were somewhat more African American participants in the revalidation sample (26.9 percent) than the construction validation sample (19.0) percent. Additionally, the construction validation represented more white participants (71.9 percent) than the revalidation sample (66.0 percent). The difference was not significant, however. See Appendix F and H for the revised WRNA and WRNA-T interviews.

5. Selection of a final risk/needs stand-alone and WRNA-T scales considered the results for the construction and revalidation samples (step 3). Total risk/needs scales were developed in the construction validation sample, retested on the revalidation sample, and then tested for specific sites. See Appendix G and I for the revised scoring forms.

Table 2. Frequency and Percent Distribution of Demographic and Offense-Related Background Characteristics of the Construction and Revalidation Samples.

	Construction		Revalidation		Total	
Characteristic	**N**	**Percent**	**N**	**Percent**	**N**	**Percent**
Age	N=322		N=318		N=640	
18-20 years old	18	5.6	14	4.4	32	5.0
21-30 years old	110	34.2	119	37.4	229	35.8
31-40 years old	104	32.2	90	28.3	194	30.3
41-50 years old	68	21.1	73	23.0	141	22.0
51 years and older	22	6.8	22	6.9	44	6.9
	$\bar{x}$ = 34.3 yrs.		$\bar{x}$ = 33.3 yrs.		$\bar{x}$ = 34.3 yrs.	
Race	N = 306		N=309		N=615	
Asian	2	0.7	1	0.3	3	0.5
African American	58	19.0	83	26.9	141	22.9
Hispanic/Latina	12	3.9	12	3.9	24	3.9
Native American	5	1.6	3	1.0	8	1.3
White	220	71.9	204	66.0	424	68.9
Other	9	2.9	6	1.9	15	2.4
Currently Married	N = 318		N=317		N=635	
Yes	76	23.9	72	22.7	148	23.3
Has Children Under 18	N=311		N=313		N=624	
Yes	194	62.4	187	59.7	381	61.1
Employment	N=318		N=316		N=634	
Employed (full or part-time, child care, student, or disabled)	205	64.5	220	69.6	425	67.1
Not employed	113	35.5	96	30.4	209	33.0
H.S. Grad or GED	N=317		N=316		N=633	
Yes	198	62.5	208	65.8	406	64.1
Current Offense	N = 312		N=313		N=625	
Violent	106	34.0	114	36.4	220	35.2
Property	97	31.1	93	29.7	190	30.4
Drug	84	29.9	84	26.8	168	26.9
Public Order	6	1.9	2	0.6	8	1.3
Other	19	6.1	20	6.4	39	6.2

Table Continues.

Table 2. Frequency and Percent Distribution of Demographic and Offense-Related Background Characteristics of the Construction and Revalidation Samples, continued.

	Construction		Revalidation		Total	
Characteristic	**N**	**Percent**	**N**	**Percent**	**N**	**Percent**
Prior Felonies[a]	N = 244		N=240		N=484	
Yes	107	43.8	113	47.1	220	45.4
Prior Incarcerations[a]	N = 243		N=240		N=483	
Yes	95	39.1	99	41.3	194	40.2

[a] Data for Rhode Island were not available on these measures.

Data analysis employed bivariate correlations (Pearson's r) and analysis of receiver operating characteristics (AUC). AUC measures were examined for cumulative scales but not the individual risk/need domains. Psychometric properties of the new scales involved factor analysis (principal component extraction and varimax rotation) and alpha reliability measures. Results for factor analysis are not shown in this report but are available from the lead author.

Offense-Related Outcome Measures

With the exception of women who were released early, participants were followed-up for 12 months, and results were reported at a 6 and 12 month intervals. Because the cumulative WRNA risk scale was designed to predict prison misconducts, the follow-up measures were: a) any prison misconduct (yes or no); b) number of prison misconducts; c) any serious misconduct (yes or no); d) number of serious misconducts; e) any aggressive misconduct (yes or no); and f) number of aggressive misconducts. Distributions on these measures are shown by state in Table 3.

It was important to examine serious and aggressive misconducts distinct from all misconducts, because they are more indicative of institutional behaviors that more clearly reflect inmates' culpability. In contrast, the more global, ANY MISCONDUCT measure includes many citations that may have emanated from misunderstandings, poor inmate-staff relationships, poor inmate management skills, and punitive institutional cultures. These more general measures of institutional misconduct capture a number of minor infractions (e.g., being out of place) that can reflect the behaviors of *both* staff and inmates. Such measures, in research parlance, are known to have more "noise" and thereby attenuate the magnitude of statistical correlations (Hewitt, Poole, & Regoli, 1984; Van Voorhis, 1994).

In the present study, serious infractions included behaviors such as drug-related behaviors, sexual contact, various forms of contraband, gambling, refusing to accept institutional assignments, manufacturing or possessing a weapon, stealing property, tampering with firearms or locks, as well as the subset of items that were also captured by the prison aggression measures. Institutional aggression measures pertaining to women typically tapped less lethal forms of aggression such as fighting and threatening behaviors. To illustrate, the most serious aggressive incidents consisted of 1 assault in Rhode Island, 4 minor assaults in Missouri, and 3 instances of "causing or attempting to cause physical harm to another" in Ohio. Thus, the aggression measures shown in Table 3 are impacted by low base rates and the fact that the majority of aggressive infractions involved fighting and threatening behaviors rather than assaults.

Table 3 also shows considerable differences in the prevalence rates of disciplinary infractions across sites. This was not unexpected. The distinctions reflect different staff and

institutional cultures, staff inmate management skills, and the appropriateness of a facility's infraction codes for women offenders.

Limited base rates and short follow-up time frames pose the risk of attenuating research findings. Low base rates affected the aggressive misconduct measures and all of the Ohio 6-month follow-up measures. Additionally, the literature in corrections typically recommends follow-up periods of two years (Maltz, 1984). These two challenges are endemic to studies of female inmates. Female inmates commit far fewer and less serious acts of aggression than men (Hardyman & Van Voorhis, 2004), and in many states they serve much shorter sentences. Extending the follow-up time frame for these participants may not have successfully addressed these challenges, because a substantial number of the inmates would not still be incarcerated at the 24 month time period.

Table 3. Follow-up Measures by Time Frame and Site.

Site	All Misconducts			Serious Misconducts			Aggressive Misconducts		
	N	%	$\bar{X}$	N	%	$\bar{X}$	N	%	$\bar{X}$
6 Month Follow-up									
Missouri (N=98)	67	68.4	1.29	28	28.6	0.42	2	2.0	0.04
Ohio (N=386)	64	16.6	0.24	54	14.0	0.18	30	7.8	0.10
Rhode Island-WRNA-T (N=156)	65	41.7	0.74	39	25.0	0.37	7	4.5	0.06
12 Month Follow-up									
Missouri (N=53)	40	75.5	2.87	23	43.4	0.87	3	5.7	0.11
Ohio (N=347)	93	26.8	0.52	80	23.1	0.36	51	14.7	0.20
Rhode Island-WRNA-T (N=69)	36	52.2	1.28	25	36.2	0.73	8	11.6	0.20

RESULTS

Revalidation of Original Institutional WRNA

Risk/Need Domains: Correlations between the 2008 WRNA risk/needs domains and prison misconducts are shown in Table 4, for the combined Ohio, Missouri, and Rhode Island samples. A number of the patterns found in earlier studies appear in this sample as well. First, there are several domains that appear to be more predictive in community settings than in institutional settings (see Van Voorhis et al., 2010; Van Voorhis et al., 2012; Van Voorhis et al., 2013). These include educational assets, family issues, and parenting issues. Across most of the WRNA studies, these appear to be far more potent predictors in community settings, because they are more approximate to the participants' lives, acting as immediate stressors or sources of resilience.

Another pattern finds that static criminal history items, which in most states are the central variables of prison custody classification systems, are less predictive or equally predictive to domains that described troubled women, e.g. anger, depression, recent substance abuse, and child abuse. Thus, needs are highly predictive of prison misconducts.

Finally, results for longer follow-up windows are stronger than those for short (6 month) follow-up windows. This is because base rates improve with time. For this reason, two to three years is the scientific standard for criminal justice prediction research.

The strongest predictors for the sample as a whole consisted of criminal history, anger, recent substance abuse, and depression. It is noteworthy, however, that correlations seldom surpassed .20. This observation may reflect baseline problems and limited follow-up time frames. At the same time, it should be recognized that the correlations are considerably higher when the data are disaggregated by site (see Tables 5 through 7).

Research staff, including interviewers, speculated that the limited findings on relationship issues and family issues were the result of structural problems with the interview itself. These limitations will be discussed below.

Risk/Domains-Specific Sites: The predictive validity of the need domains for each site are shown in Tables 5 through 7 below.

The Missouri findings were especially strong for several of the risk/need domains such as antisocial friends, anger, psychosis (symptoms), child abuse, PTSD, and relationship dysfunction. Fewer significant findings appear for the 12 month follow-up window because the sample size was small (N=53) and statistical tests were considerably less powerful. With these less powerful tests, correlations needed to surpass .20 in order to reach statistical significance.

Among the Missouri participants, criminal history, educational needs, and the interview measure of parental difficulties all predicted in the wrong direction. These findings may implicate the instability of small samples. In a separate report prepared for Missouri officials, analyses determined that educational and criminal history scales were attenuated by interviewers' failure or inability to verify items with official records (Van Voorhis et al., 2012). Just the same the findings caused researchers to encourage Missouri officials to view these results with caution. As will be seen later, however, it was possible to revise the criminal history and other scales to improve upon these findings.

With a larger sample, tests for the Ohio site produced more significant findings than those seen for Missouri. Table 6 shows that the strongest correlates were observed for measures of criminal history and anger. Other findings, while significant, were somewhat more modest than those seen for the Missouri and the Rhode Island samples. It is not clear why this was the case, but we noted earlier that this was a research sample and female inmates in Ohio had been

Table 4. Bivariate Relationship between Original WRNA Scales and Prison Misconducts, All Sites.

	6 month (N=484 & 640)[a]						12 month (N=400 & 470)[b]					
	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N
Interview												
Criminal history	.10**		.11***	.07**	.12***	.09**	.17***	.11**	.18***	.16***	.19***	.16***
Attitudes	.06*		.09**	.06*		.06*						
Educational needs	.08*						.07*	.08*				
Educational assets[c]	-.07**	-.11***		-.07**		-.05*	-.06*	-.11***	-.07*	-.09**	-.08**	-.10**
Employment/financial[c]	.12***	.12***	.08**	.08**	.07**	.06*	.10**	.15***	.13***	.14***	.12***	.12***
Antisocial friends	.10**		.12***	.09**	.09**	.06*	.11***	.08**	.14***	.08**	.14***	.12***
Anger[c]	.16***	.14***	.17***	.17***	.28***	.26***	.24***	.19***	.25***	.24***	.32***	.32***
Mental health hist.[c]	.09***		.06*				.09**					
Depression (symptoms)[c]	.09***	.08**	.10***	.09***	.12***	.11***	.07*	.08**	.08**	.10**	.15***	.15***
Psychosis (symptoms)[c]									.07*	.10***	.09**	.07*
Child abuse (interview) [c]	.09***	.07**	.12***	.10***			.08*	.09*	.08*	.16***	.11**	.08*
Adult abuse (interview) [c]												
Sex abuse (adult or child) [c]	.06*		.08**	.07**							.06*	
Physical abuse [c]	.07**		.08**							.09*		
PTSD[c,d]	.08**	.07*	.12***	.10***	.07**	.07*	.11**	.08**	.10**	.08**	.11**	.09**
Substance abuse history							.08**		.08*			
Substance abuse (recent)[c]	.18***	.17***	.11***	.09**			.19***	.14***	.16***	.08*		
Parental involvement (strength)[c,e]	.07*	.08*			.07*							
Parental difficulties (all)		-.06*	-.06*	-.07**					-.07*	-.08**	-.07*	-.07*
Relationship satisfaction (strength) [c]												
Family conflict [c]												
Family support (strength) [c]												
Survey												
Relationship dysfunction [c]			.08*									
Self-efficacy (strength) [c]	-.06*						-.06*				-.09**	
Child abuse [c]	.07**		.13***	.09**	.05*	.05*			.07*	.08**	.11***	.10***
Adult abuse [c]	.06*											
Parental stress [c]											-.06*	

***p≤.01, **p≤.05, *p≤.10

[a]Participants were in the 6 month window for at least 3 of the 6 months. N for scales appearing on the full WRNA only=484; N for scales appearing on the trailer and the full WRNA=640.

[b]Participants were in the 12 month window for at least 9 of the 12 months. N for scales appearing on the full WRNA only=400; N for scales appearing on the trailer and the full WRNA=470.

[c] Indicates a scale appearing on both the full WRNA and the trailer (N=640 at 6 mos.)(N=470 at 12 months).

[d] Questions were not asked by all interviewers (N=535 at 6 mo.)(N=414 at 12 months).

[e] Scale pertains to mothers (N=381 at 6 mo; N=273 at 12 mo.)

Table 5. Bivariate Relationship between Original WRNA Scales and Prison Misconducts, Missouri.

	6 month (N=98)[a]						12 month (N=53)[b]					
	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N
Interview												
Criminal history	-.15*	-.20**									.20*	
Attitudes	.16*											
Educational needs			-.30***	-.28***	-.13*	-.13*			-.24**	-.23**	-.21**	-.19*
Educational assets	-.16*	-.13*										
Employment/financial												
Antisocial friends	.28***		.25***	.14*			.26**					
Anger	.18**	.14*		.14*	.25***	.25***		.21*			.22*	.30**
Mental health hist.	.14*								-.28*	-.30***		
Depression (symptoms)												
Psychosis (symptoms)		.20**	.21**					.42***				
Child abuse (interview)	.16*	.13*	.34***	.29***	.14*	.14*				.23**		
Adult abuse (interview)												
Sex abuse (adult or child)	.18**		.26***	.17**							.20*	
Physical abuse	.10**		.23**	.15*								
PTSD	.21**	.15*	.17**				.33***	.33***				
Substance history	.14*		.14*									
Substance abuse (recent)	.17**	.15*	.17**									
Parental involvement (strength)[c]							.31**			.27*		
Parental difficulties (all)[c]		-20**			-.16*	-.16*					-.22*	
Relationship satisfaction (strength)							.22*					
Family conflict												
Family support (strength)		.17**										
Survey												
Relationship dysfunction	.17**		.34***	.19**								
Self-efficacy (strength)										.18*		
Child abuse	.19**	.16**	.35***	.25***								
Adult abuse	.14*											
Parental stress	.15*											

***p≤.01, **p≤.05, *p≤.10

[a]Participants were in the 6 month window for at least 3 of the 6 months..

[b]Participants were in the 12 month window for at least 9 of the 12 months.

[c] Interview parenting questions were not asked by all interviewers (N=68).

asked to participate in an excessive number of other research studies during the same time frame. In contrast, the Rhode Island participants, and to a lesser extent the Missouri participants, were engaged in a course of treatment that could impact their future. They were interviewed by their case managers rather than researchers. The Rhode Island participants, especially, may have been more motivated to participate in the interview.

Rhode Island tested the WRNA-Trailer as a supplement to the LSI-R. Thus, gender-neutral risk/need domains were obtained from the LSI-R interviews, administered to the same participants. Table 7 indicates that the LSI-R domains were strongly associated with aggressive misconducts by the 12 month follow-up point. A number of the original gender-responsive domains, measured by the WRNA-T, were also associated with prison misconducts. These included anger, depression, psychosis, adult abuse, sexual abuse, PTSD, and self-efficacy.

Total Risk Scale: The original total risk scale represented the total score of all risk/need domains found to be predictive in the 2008 construction validation study. This scale summed scores for criminal history, antisocial attitudes, anger, psychosis, family conflict and collapsed measures of mental health history, depression, child abuse, substance abuse history, and relationship dysfunction. The original scale also subtracted a strength, family support (collapsed), from the total score. Ideally, correlations should surpass .27, and AUC values should surpass .70. Such results were seen for some of the Ohio tests but not for the Missouri analysis.

WRNA-T: The Trailer for the Women's Risk/Needs Assessment (WRNA-T) was comprised of gender-responsive items, i.e., anger, psychosis, and collapsed measures of mental health history, depression, family support and relationship dysfunction. For purposes of risk assessment, the trailer was not intended to be used alone but rather to be added to the LSI-R, or

Table 6. Bivariate Relationship between Original WRNA Scales and Prison Misconducts, Ohio.

	6 month (N=386)[a]						12 month (N=347)[b]					
	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N
Interview												
Criminal history	.16***	.12***	.15***	.13***	.13***	.13***	.20***	.20***	.21***	.21***	.20***	.17***
Attitudes	.12***	.08***	.12***	.11**						.08*		
Educational needs									.07*			
Educational assets								-.10**	-.08*	-.09**	-10**	-.11**
Employment/financial		.09**	.09**	.10**	.10**	.08**	.08*	.15***	.13***	.15***	.14***	.16***
Antisocial friends	.08**	.08*	.09**	.07*	.10**	.08*	.11***	.13***	.15***	.13***	.15***	.14***
Anger	.30***	.26***	.27***	.26***	.26***	.25***	.35***	.32***	.35***	.31***	.32***	.32***
Mental health hist.	.12***		.09**				.09**		.07*			.11*
Depression (symptoms)	.10**		.09*		.13***	.09**	.08*		.09*		.12**	.11**
Psychosis (symptoms)					.07*						.10**	
Child abuse (interview)	.10**		.10**	.07*			.08*		.09**	.07*	.07*	
Adult abuse (interview)												
Sex abuse (adult or child)												
Physical abuse	.08*											
PTSD			.07*				.08*		.08*		.07*	
Substance history							.07*					
Substance abuse (recent)							.11**		.13***		.11**	.07*
Parental involvement (strength)[c]												
Parental difficulties (all)											-.08*	-.07*
Relationship satisfaction (strength)					-.10*							
Family conflict	.09**		.07*									
Family support (strength)												
Survey												
Relationship dysfunction	.07*		.09**	.07*								
Self-efficacy (strength)									-.07*		-.10**	
Child abuse	.12***	.07*	.11***	.07*	.10*	.09*	.10**	.08*	.12**	.09**	.09**	.09**
Adult abuse	.08*											
Parental stress											-.08*	

***p≤.01, **p≤.05, *p≤.10

[c]Participants were in the 6 month window for at least 3 of the 6 months.

[b]Participants were in the 12 month window for at least 9 of the 12 months.

[c] Interview parental involvement scale pertains to mothers (N=231 at 6 mo.)(N=205 at 12 mo.).

Table 7. Bivariate Relationship between LSI-R and Original WRNA Scales and Prison Misconducts, Rhode Island.

	6 month (N=154)[a]						12 month (N=69)[b]					
	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N
LSI-R Gender-Neutral Scales Interview												
Criminal history												
Education/employment	.11*	.20***			.17**	.14**	.25**	.35***	.17*	.28***	.19*	.23**
Financial	.15**	.15**					.16*	.21**		.18*	.17*	
Family/marital		.14**									.18*	.19*
Accommodation	.11*	.12*								.17*	.18*	.17*
Leisure/recreation	.18***	.18***	.11*			.11*			.21**	.27***	.24**	.23**
Companions					.11*				.16*	.17*	.34***	.23**
Alcohol/drugs						.11*					.15*	.25**
Emotional/personal	.22***	.22***		.13*	.14**	.12*					.19*	
Attitudes/orientation		.21***		.13*	.16**	.24***			.28***	.24**	.36***	.43***
Total LSI-R Scale	.15**	.21***			.15**	.15**				.25**	.32***	.34***
WRNA-Interview Scales												
Educational assets					-.15**	-.15**						-.16*
Employment/financial												
Anger	.17**	.25***		.17**	.34***	.30***	.18*	.26**	.22**	.28***	.38***	.44**
Mental health hist.	.12*	.17**		.13**								
Depression (symptoms)	.20***	.28***	.21***	.22***	.13*	.18***	.21**	.37***	.29***	.36***	.37***	.43***
Psychosis (symptoms)							.17*	.23			.32***	.19*
Child abuse (interview)	.14**	.14*								.16*		
Adult abuse (interview)			.12*	.13**			.21**		.25**	.19*	.19*	
Sex abuse (adult or child)	.14**	.14**		.12*				.17*	.16*	.22**	.28***	.16*
Physical abuse					.10*							
PTSD	.22*	.28**	.21*	.26**		.21*		.43**		.40*	.56***	.49**
Parental involvement (strength)[c]	.15*											
Parental difficulties (all)[c]												
Relationship satisfaction (strength)	.13*	.20***	.12*	.16**								
Family conflict					.10*							
Family support (strength)			-.11*									

Table Continues.

Table 7. Bivariate Relationship between LSI-R and Original WRNA Scales and Prison Misconducts, Rhode Island, Continued.

	6 month (N=154)[a]						12 month (N=69)[b]					
	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N
WRNA Survey Scales												
Relationship dysfunction												
Self-efficacy (strength)	-.19***	-.19***		-.11*	-.13**	-.14**	-.23**		-.20**			
Child abuse											.17*	.16*.
Adult abuse												
Parental stress							.25**		.16*			

***p≤.01
**p≤.05
*p≤.10

[c]Participants were in the 6 month window for at least 3 of the 6 months..

[b]Participants were in the 12 month window for at least 9 of the 12 months.

[c] Interview parenting questions were not asked by all interviewers (N=81 at 6 mo;N=31 at 12 mo).

Table 8. Bivariate Relationship between Original WRNA Cumulative Scales and Prison Misconducts

	6 month						12 month					
	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N
Full Original WRNA Assessment (Missouri & Ohio)(N=484 at 6 mo)(N=400 at 12 mo.)												
WRNA-Full Scale	.16***	.06	.21***	.14***	.18***	.15***	.20***	.10**	.20****	.17***	.23***	.20***
AUC	.61		.66		.79		.63		.64			
Levels	.17***		.18***	.12***	.14***	.13***	.20***	.10**	.17***	.13***	.19***	.18***
AUC	.61		.64		.66		.62		.61			
Missouri (N=98 at 6 mo., N=53 at 12 mo.)												
WRNA-Full Scale			.16*								.20**	
AUC	.55		.58		.62		.46		.41		.70	
Levels	.15*								-.23**			
AUC	.59		.57		.45		.49		.38		.56	
Ohio (N=386 at 6 mo., N=347 at 12 mo.)												
WRNA-Full Scale	.27***	.18***	.25***	.19***	.18***	.15***	.27***	.23***	.28***	.24***	.23***	.21***
AUC	.72		.72		.70		.69		.70		.69	
Levels	.25***	.19***	.22***	.19***	.16***	.16***	.26***	.23***	.25***	.22**	.21***	.21**
AUC	.69		.68		.68		.67		.67		.67	

***p≤.01
**p≤.05
*p≤.10

an alternative gender-neutral risk/needs assessment. The WRNA-T could be tested on its own in all sites. However, its contribution to an existing gender-neutral assessment could only be tested in Rhode Island, where the LSI-R was used as the gender-neutral assessment.

The gender-responsive variables by themselves offered statistically significant predictions of prison misconducts for all of the sites (see Table 9). However, results were stronger for the Rhode Island and Ohio sites than for the Missouri site. Correlations between the WRNA-T and outcome measures are shown in Table 9 for purposes of illustration. We do not apply a standard of r=.27 and AUC of .70 to the WRNA-T, because, by itself, the WRNA-T is not intended to serve as a risk assessment.

When used as a supplement to the LSI-R in Rhode Island, the predictive merits of the WRNA-T were stronger than those for the LSI-R. As a result, the predictive validity of the combined WRNA-T and LSI-R was lower than that for the WRNA-T alone, but considerably higher than correlations for the LSI-R, alone. Moreover, on most outcome measures, the variation attributable to the WRNA-T alone (partial correlation) was significant. In fact, for the 12 month window, partial correlations with the aggression measures were quite strong.

Revisions to the WRNA

With the exception of two scales (self-efficacy and PTSD) most of the original 2008 WRNA scales were developed by UC researchers and members of the Missouri Women's Issues Committee and tested on one rather small, prison sample (N=272). With validity and scale construction test limited to one sample, it was imperative to examine ways to refine the WRNA scales.

Table 9. Bivariate Relationship between Original WRNA-Trailer Cumulative Scales and Prison Misconducts

	6 month						12 month					
	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N
Full Original WRNA Assessment (Missoui, Ohio, & Rhode Island)(N=640 at 6 mo)(N=470 at 12 mo.)												
WRNA-Trailer All Sites	.13***	.10***	.17***	.14***	.19***	.17***	.14***	.13***	.17***	.17***	..25***	.24***
Missouri (N=98 at 6 mo.)(N=53 at 12 mo.)												
WRNA-Trailer	.17**		.27***	.21***	.16*	.16*		.24**		.18*	.20**	.18*
Ohio (N=386 at 6 mo.)(N=347 at 12 mo.)												
WRNA-Trailer	.25***	.18***	.23***	.18***	.19***	.17***	.25***	.21***	.25***	.22***	.24***	.23***
Rhode Island Intake (N=154 at 6mo)(N=69 at 12 mo.)												
LSI-R Total Scale	.15**	.21***			.15**	.15**		.28***		.25**	.32***	.34***
WRNA-Trailer	.17**	.25***	.13**	.16**	.17**	.16**		.34***	.22**	.34***	.40***	.42***
LSI-R + WRNA-Trailer	.17**	.24***		.10*	.18**	.17**		.32***	.17*	.30***	.37***	.39***
AUC	.59				.77				.60		.80	
Partial corr.	.12*	.17**	.14**	.14**	.11*			.25**	.18*	.26**	.29***	.32***

***p≤.01
**p≤.05
*p≤.10

The methodology for improving scales involved a series of item analyses conducted on a new construction validation sample systematically selected from the current research sample. Scales were then retested on a revalidation sample, also systematically selected from the current research participants. As noted earlier, the two samples were drawn through a systematic random sampling procedure, and there was no overlap of participants between the construction and revalidation sample. Use of the two samples was intended to bolster confidence in the validity of the revised tool by securing revalidation findings through the present study.

Tests of the revised scales are shown in Tables 10 and 11, below for the construction and revalidation scales, respectively. The discussion of scale revisions and psychometric findings follows. A table showing measures of scale internal consistencies appears in Appendix E.

CRIMINAL HISTORY: Problems with the original criminal history scale were detected during the original construction validation research. It was assumed that the scale would be amended as part of the present revalidation study. To assist with this effort, research personnel in the Missouri Department of Corrections suggested six additional questions for the scale. The items that contributed to the predictive validity of the scale were:

1. Was your last conviction within the past three years?
2. Age at intake
   18-34 = 2
   35-45 = 1
   46+ = 0

In addition to these changes, validity was improved considerably by using official accounts of prior felonies and number of previous incarcerations. The practice of asking offenders for this information should be discouraged. As shown in Table 10 and 11, these changes improved the scale for the construction sample, but results for the revalidation sample were significant for the 12 month group, but not the 6 month group. Improvements were nevertheless evident for both the Ohio and Missouri samples, so the changes were accepted. In institutional settings, the validity of this scale is attenuated by the need to include violent offenses. One item indicates whether or not the present offense was violent and another indicates the presence of prior convictions for violent offenses. Often these items are not predictive for female inmates. However, it is unlikely that sites would opt for excluding this information from an institutional assessment tool. Alpha for the scale was somewhat low (.55) but that is not unusual for criminal history scales.

ATTITUDES: The attitudes scale was found to be significantly related to prison misconducts in the Ohio sample, but only one of the Missouri tests was significant. In this study, the scale alpha was somewhat low (.67), but it was considerably higher for the pre-release and probation studies. Item analysis revealed that many of the individual items were not related to prison misconducts. As a result, there were no opportunities to improve this scale. It should be noted that the LSI-R Attitudes/orientation scale *was* correlated to prison misconducts in the Rhode Island sample. The discrepancy may suggest that the WRNA interviewers were not adept at hearing criminal thinking patterns. Improvements to staff training protocols may be needed. However, in other UC WRNA studies, the anger and self-efficacy scales typically offered far better contributions to the predictive validity of the total scale than the WRNA attitudes scale. As a result, the attitudes scale is included in the needs section of the assessment (Part IV) and not in the risk calculation.

EDUCATIONAL SCALES: For the sample as a whole, the <u>educational needs</u> scale had limited effects on prison misconducts. One variable, "have you attended any special education classes" appeared to detract from the scale; it was negatively related to outcomes. However, its removal did not improve the scale. Alpha for the original scale was .66, but was higher in the probation and prerelease studies.

The scale for <u>educational strengths</u> showed several modest correlations with prison outcomes. However, it too could not be improved and alpha was also somewhat low, .65. The results conformed to a pattern affecting a number of scales that were more predictive in community samples than institutional. Indeed, educational assets are found to be a source of resilience in both the 2008 and 2013 probation settings (Van Voorhis et al., 2010; Van Voorhis et al., 2013). For the prison instrument, these scales are not included in the cumulative final risk needs scales, but instead appear in Part IV.

EMPLOYMENT/FINANCIAL: The employment/financial scale predicted prison outcomes in the Ohio samples, but not the Missouri or the Rhode Island samples. Alpha for the original scale was .71. A number of items were tested to improve the predictive validity of the scales. Some of these were designed to provide an opportunity to tap economic dimensions associated with poverty rather than middleclass life styles. As a result, we tested a number of items that were suggested by practitioners and administrators in the study sites. They did not improve the predictive validity over what was seen in Tables 4 through 7. As a result, no changes were made to this scale and it is not included in the risk scale.

ANTISOCIAL FRIENDS: The Antisocial Friends scale offered modest correlations with prison misconducts. Item analysis revealed that one item, "prior to your arrest, did you have some friends who seemed supportive of you?", was detracting from the validity of the scale. Its removal slightly improved the predictive validity of the scale as well as its alpha reliability (.77). The scale is included in Part I.

ANGER: Across samples, the WRNA Anger Scale tended to be a more potent predictor of offense-related outcomes than the other cognitive variables, e.g., attitudes (antisocial thinking). This could be attributable to the fact that the antisocial thinking scale was more subjective while the anger scale was behavioral. However, earlier focus groups with women offenders informed us that anger ("rage) was a significant factor in their offending. Strong correlations were observed in this study as well, and alpha was sufficient .73. No revisions were made to this scale. It is included in the institutional risk scale, Part I, of the assessment.

HISTORY OF MENTAL ILLNESS: The History of Mental Illness scale was seldom predictive of prison misconducts. In other writings, we have speculated that this was attributable to the fact that different mental health problems were combined into a single scale, when in fact some mental health diagnosis were predictive and others were not (Van Voorhis et al., 2010). Combining items therefore allowed predictive items to be cancelled-out by non-predictive items. However, in this case, none of the single items were significantly related to outcomes. It is also possible that a change in mental health status, such as that which might accompany successful treatment, might reduce the predictive merits of this scale. The scale, however, is retained for case management purposes in a separate section of the assessment (Part IV). Alpha for this scale was .78.

CURRENT SYMPTOMS OF MENTAL ILLNESS: In most samples, current symptoms of mental illness tended to be more predictive than the static History of Mental Illness Scale.

> **DEPRESSION:** The original WRNA Depression Scale was weakly correlated with some measures of prison misconducts for the sample as a whole. However, it was associated with misconducts in the Rhode Island and Ohio samples. Two new items were tested, but they did not make any improvements to the scale. The scale was collapsed into low (0), medium (1-4) and high (5) for use in the final risk scale. These are the same recodes as used on the pre-release instrument. Results for the collapsed scale were significantly related to outcomes in both the validation and the revalidation samples. Alpha for the scale was .77. The collapsed scale is included in the prison risk scale.
>
> **PSYCHOTIC SYMPTOMS**: This scale was comprised to two items which showed modest predictions to some outcomes in all samples. There were no test variables to make any amendments to the scale. The inter item correlation (alpha was inappropriate) was $r=.34, p\leq .001$.

ABUSE-INTERVIEW SCALES: For the combined samples, child, sexual, and physical abuse were significantly correlated to prison misconducts. However, relationships were stronger for Missouri and Rhode Island samples than they were among the Ohio participants. Even so, the child abuse scale was found to be related to outcomes in all of the sites as well as in all of the prison sites participating in the earlier 2008 study. It is important to note that these items show interviewer effects, where results were stronger for some interviewers than others. Such findings will require changes to training

protocols which stress the importance of good interviewer-participant rapport and of giving clear examples of abusive behaviors prior to asking the four questions.

It is important to note that the interview items, nevertheless achieved predictions that in most cases were as strong as those contained in the survey at the end of the assessment. As will be discussed below, these findings furnished the rationale for omitting the two survey abuse sections. The interview child abuse scale appears on the risk scale in Part I; adult abuse, physical abuse, and sexual abuse scales are listed in Part IV for purposes of guiding case management.

PTSD: Four interview items were based upon the Veteran's Administration's Post Traumatic Stress Disorder Scale. The cumulative scale was associated with misconducts at all sites. Correlations were especially strong for Missouri and Rhode Island but rather weak for the Ohio sites. However, early interviews in Rhode Island did not contain the items. As a result the scale was marred by missing data. In this study, there was no way to repair this situation, because too many cases were affected. A decision was made to keep the scale on the assessment for further study, because when the data were available, the scale was highly predictive. It was not included in the risk/needs scale for the present version of the WRNA. Alpha was .78.

SUBSTANCE ABUSE: Two substance abuse scales were created for the WRNA--substance abuse history and recent substance abuse. Results were much stronger for the recent substance abuse scale than the substance abuse history scale. The present study examined the addition of two questions, one for each scale.

> **SUBSTANCE ABUSE HISTORY**: Item analysis revealed that very few items on this scale were related to prison misconducts. Moreover the addition of a test item "Does the offender have substance abuse-related offenses on record," failed to improve its predictive validity. The scale appears in Part IV of the assessment. Alpha for the revised scale was .79.
>
> **SUBSTANCE ABUSE CURRENT**: This scale was often strongly associated with offense-related outcomes. Just the same, the study afforded an opportunity to improve the scale. A test item, "Do you currently have any feelings that you need to use drugs first thing in the morning," strengthened the correlation in both the validation and the revalidation samples. Alpha for the new scale was somewhat low (alpha = .63), but it is known to be considerably higher in other studies (.76 for probationers and .71 for pre-release participants).

FAMILY OF ORIGIN SCALES: Two scales were created, one measuring Family Conflict and another Family Support. As can be seen in Tables 4 through 7, neither scale appeared to be correlated with misconducts. A number of difficulties during the interview process may explain these findings. Most importantly both scales are skewed

to showing extremely favorable evaluations of families. In addition interviewers reported that many participants had difficulties conceptualizing siblings and parents in contexts of blended families, separated families, parent figures, etc. Finally, some respondents substituted a number of concepts of "family" for questions that only pertained to family of origin or adoptive family. These difficulties could not be corrected in the present study. However, the revised interview is restructured to capture a more clear definition of family of origin and will be tested in future studies.

FAMILY CONFLICT: One test items was added to this scale to make modest improvements to predictive validity: "Do your parents or any siblings tend to be critical of you when they communicate with you?"

The item was omitted by a number of the Rhode Island interviewers.[9] As a result, the findings shown in Tables 10 and 11 pertain only to Ohio and Missouri. Alpha for this scale was low (.40), and has been in all of the studies, indicating that the scale is picking up diverse dimensions of family conflict. The Family Conflict scale appears in Part IV of the assessment.

FAMILY SUPPORT: The family support variable, noted to have favorable findings in other studies, was not observed to be negatively related to prison misconducts in these samples. As with the family conflict scale, however, results were quite skewed to favorable (supportive) ratings of families, and variability on the scale was somewhat limited. No test variables were entered into the study and collapsing the scale to values similar to those used in other studies made no improvements to its predictive validity. Nevertheless the scale did improve the overall predictive validity of the institutional risk scale. The collapsed scale appears in Part II and is subtracted from the total risk score. Alpha for this scale was .75.

PARENTING SCALES: Parenting scales were predictive in several of the community samples. Three were tested, one pertaining to parental involvement and two to parental stress. The purpose for testing two versions of parental stress/difficulties was to determine whether it would be possible to omit one of the scales to assist efforts to shorten the interview process. In doing so, it was necessary to compare the predictive validity of the two scales.

PARENTAL INVOLVEMENT: Parental Involvement was a source of resilience in community settings, especially probation settings. In institutional

[9] For much of the study period, Rhode Island used an interview that did not have the test questions, because the Rhode Island studies began almost a year prior to the Ohio and Missouri studies.

settings, however, it was not. In fact, we occasionally saw positive rather than negative associations with prison misconducts. Item analysis revealed no opportunities to improve the predictive validity of the scale. It was retained in Part IV of the assessment to assist prerelease planning. The scale alpha was .72.

PARENTAL DIFFICULTIES: The attempt to create an interview scale to substitute for the survey scale was not successful. Analyses of probation and parole samples revealed that the survey scale (below) was superior. Therefore, the revised interview retains only four of the eight questions appearing in the original interview. These are retained for purposes of case management.

PARENTAL STRESS: As with the other parenting scales, the survey parental stress scale was seldom correlated with prison outcomes. It is retained in the needs assessment section of the assessment (Part IV), because it has been observed in community samples to be a risk factor. As such, the scale is needed to aid in referrals to parenting classes. Alpha was .83.

There were some contradictions between the interviewer's indications of whether the woman has children under 18, and the women's indication on the survey. As a result, the scale was keyed to the interviewer's indication of whether the woman had children. Modifications will be made to the training protocol, to recommend that the interviewer determine that the woman has had at a period of ongoing contact with any children who are 18 or younger at the time of the interview. Correcting for this must be done during the scoring/research process. Therefore, the questions do not pertain to women who have never had a period of ongoing contact with any children who would have been under 18 at the time of the interview. These women and other non-parents are scored as zero on this scale. The scoring steps are as follows:

1. Total the scale items for women who have children under 18 with whom they have had contacts with.
2. For that group of women, replace any missing cases at the median.
3. Once the first two scoring steps have been completed, non-parents are entered into the scale as 0.

INTIMATE RELATIONSHIP SCALES: The original survey scale for relationship dysfunction was correlated with prison misconducts among Missouri participants, modestly correlated among Ohio participants and not correlated with prison misconduct for the Rhode Island sample. Moreover, obtaining participants' responses to these questions incurred a number of difficulties. First, interviewers reported that women were

very guarded in their discussions of significant others. As a result, the relationship items had more missing values than other items. Second, researchers observed that interviewers sometimes interjected their own evaluations of whether the woman was actually involved in an appropriate relationship. Interviewers would then alter survey results accordingly. Third, a number of items designed to tap dimensions of supportive and satisfying relationships produced findings that some would find counterintuitive. For example, women involved in satisfying relationships tended to be somewhat *more* likely to be involved in prison misconducts than those not involved in satisfying relationships.

RELATIONSHIP DYSFUNCTION: It was possible, however, to construct a smaller scale of relationship dysfunction which contained 4 rather than 6 items.

The items comprising this scale are as follows:

1. Do you find yourself more likely to get in trouble with the law when you are in a relationship than when you are not in a relationship?
2. Do you tend to get so focused on your partner that you neglect other relationships and responsibilities?
3. Have partners been able to convince you to get involved in criminal behavior?
4. Do you feel okay about yourself when you are not in a relationship? Or if in a relationship: Would you feel okay about yourself if you were not in a relationship?

Alpha for this scale was low (.60) but improves to .72 without the fourth item, "Do you feel okay about yourself when you are not in a relationship?" This may reflect the fact that the fourth item was collected by interview rather than by survey. The revised assessment will move the item to the survey portion, and that may improve the internal consistency of the scale in future tests. An examination of table 10 and 11 shows that the new scale showed improved predictive validity within the construction validation sample at 6 and 12 months, but results for the revalidation sample were not significant. The scale does, however, improve the overall predictive validity of the scale in all of the state samples. For that reason, a decision was made to include the scale in Part I of the assessment.

RELATIONSHIP SATISFACTION: The first three items on the original survey spoke to a sense of satisfaction in intimate relationships.

1. In general would you describe these relationships as supportive and satisfying?
2. Do you get into relationships that are painful for you?

3. Have significant others loved and appreciated you for who you are?

When combined, these items formed a scale (alpha=.72). The scale, however, was not found to be related to prison misconducts. To reduce the possibility of interviewer bias and improve the privacy of the questions, data for both scales will be collected in the survey portion of the assessment. The scale appears in Part IV.

RELATIONSHIP STABILITY: The few remaining interview relationship questions collect necessary information on marital status and whether or not the woman has a significant other. Together, these items did not form a scale. Consistent with the survey responses the item "Are you involved with a significant other" was positively associated with prison misconducts. The items pertaining to marital status, and the length of the marriage, however, were negatively related to prison misconducts. Married women and those involved in long term relationships were less likely to incur prison misconducts. We did not, however, include marital status in a risk scale. It appears in Part IV.

SELF EFFICACY: The self-efficacy scale was the well-established Sherer Self-Efficacy scale (Sherer, Maddux, Mercadante, Prentice-Dunn, Jacobs & Rogers, 1982)(Alpha=.89) that we did not wish to make improvements to. The scale seldom correlates with prison misconducts. However, it is an important factor in adjustment upon release and contains a number of items pertinent to key cognitive skills of decision making, problem solving and making plans for the future. It was retained in the needs assessment section (Part IV) but is not added to the prison risk scale.

ABUSE SURVEY SCALES: Survey abuse items were originally developed to provide privacy to participants who might rather indicate abuse histories on a pencil and paper survey rather than discuss these situations with an interviewer. Researchers also wished to avoid the problem of under-reporting caused by participants who did not know that specific experiences were abusive (Browne, Miller, & Maguin, 1999). That is, the survey items corresponded to criteria for determining whether or not abuse had occurred, and it was assumed that behavioral measures would better capture the experiences of the respondents. As can be seen on Tables 4 through 7, above, however, the interview was as effective as the survey at securing predictive indicators of abuse. In addition to providing minimal gains over the interview questions, the survey questions proved emotionally difficult for some respondents. These observations provided an opportunity to omit the survey abuse questions on all WRNA assessments (Pre-release, Probation, and Institutional). Doing so reduced the assessment by 32 questions.

In securing the information from the interview, however, interviewers are directed to begin the section by giving examples of behaviors considered to be abusive (see Appendix F and Appendix H).

Table 10. Bivariate Relationship between Revised WRNA Scales and Prison Misconducts, Construction Validation Sample.

	6 month (N=244/322)[a]						12 month (N=209/240)[b]					
	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N
Interview												
Criminal history	.25***	.23***	.25***	.21***	.18***	.18***	.31***	.30***	.35***	.33***	.25***	.25***
Antisocial friends	.11**		.14***	.11**	.10*	.10*	.11*		.13**		.16***	.13**
Depression (symptoms) Recoded	.11**	.09**	.11**	.09*	.14***	.14***			.09*	.10**	.18***	.19***
Substance abuse (recent)	.22***	.22***	.14***	.13**			.22***	.20***	.15***	.12**		
Relationship stability (strength)	-.13***	-.10**	-.09**	-.11**	-.13***	-12**	-.08*		-10*	-.10*	-.17***	-.16***
Family conflict[c]	.12**		.09*									
Family support (recoded)	.09*											
Survey												
Relationship dysfunction	.11**	.07*	.16***	.10**							.18***	.10*

***p≤.01
**p≤.05
*p≤.10

[c]Participants were in the 6 month window for at least 3 of the 6 months..

[b]Participants were in the 12 month window for at least 9 of the 12 months.

[c] Rhode Island cases are omitted, because interviewers could not ask test questions. (N=244 at 6 mo; N=209 at 12 mo

Table 11. Bivariate Relationship between Revised Scales and Prison Misconducts, Revalidation Sample

	6 month (N=240/318[a]						12 month (N=191/230)[b]					
	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N
Interview												
Criminal history							.17***	.14**	.24***	.21***	.17***	.11*
Antisocial friends	.09*		.11**	.09*			.13**	.13**	.17***	.12**	.12**	
Depression (symptoms) Recoded		.08*	.14***	.13***	.14***	.13***		.10*	.13**	.12**	.15***	.14**
Substance abuse (recent)	.18***	.13***	.09*				.18***	.14**	.18***			
Relationship stability (strength)	-.07*				-.08*					-.09*	-.11**	-.10*
Family conflict[c]			.16***	.13**			.12*	.15**	.12*	.20***		.10*
Family support (strength)												
Survey												
Relationship dysfunction												

***p≤.01
**p≤.05
*p≤.10

[a]Participants were in the 6 month window for at least 3 of the 6 months..

[b]Participants were in the 12 month window for at least 9 of the 12 months.

[c] Rhode Island cases are omitted, because many interviewers did not ask the test questions. (N=240 at 6 mo; N=191 at 12 mo)

In sum, test questions and item analysis resulted in changes to 8 scales, criminal history, antisocial friends, depression, recent substance abuse, relationship stability, relationship dysfunction, family conflict, and family support. In most cases, the revisions also improved the predictive validity of individual risk/needs scales in each of the research sites. Site-specific results are shown in Appendix J. Revisions also, resulted in the omission of two survey scales (adult and child abuse) and one interview scale (parental difficulties). This shortened the overall assessment process from 190 questions to 136 questions.

Unfortunately, some scales remain in need of improvements. For example, the structure and context of the family support and family conflict scales need to be improved through additional research. In addition, it is likely that the relationship scales will continue to be sample specific.

Revision of the Total Risk Scale

Revisions to the above scales also improved the predictive validity of the total risk/needs scale over results shown for the 2008 WRNA shown in Tables 8 and 9.

The risk scale for the standalone WRNA summed the risk/need factors found to be significantly related to institutional misconducts, and then subtracted strengths. The scoring form (Appendix G) adds the following risk/need scales in Part I:

Criminal history
Anger
Antisocial friends
Recent substance abuse
Depression symptoms (collapsed)
Psychotic symptoms
Child abuse
Relationship dysfunction

Family support is subtracted from this total and appears in Part II of the scoring form. Family support was not found to be predictive in most of the samples, but we believe this to be attributable to misunderstandings that interviews had with the questions. It was predictive in earlier studies, and its addition at this point does not detract from the predictive validity of the risk scale.

As can be seen in Table 12, the revised WRNA scale was highly predictive for the construction validation sample at 12 months, but somewhat less so for the revalidation sample. The criminal history score is shown as a basis of comparison, because similar static criminal history scores form the basis of custody classification systems used in both men's and women's prisons. For the construction validation sample, correlations for the static criminal history scale outperformed the WRNA in 5 of 12 tests. However, AUC values for the WRNA were higher than the criminal history scale in 5 of 6 tests. For the revalidation sample, correlations and AUC values for the WRNA surpassed those for the criminal history scale in all tests.

For the construction validation sample, correlations for the stand alone WRNA surpassed .27 and AUCs surpassed .70 on predictions of serious offenses at both the 6 and 12 month point. Similar findings were seen for predictions of aggressive misconducts at 12 months. Correlations were somewhat lower for the revalidation sample. However, acceptable results were seen for the 12 month cohort.

The revised WRNA Trailer was comprised of the subset of scales considered to be gender-responsive (see Appendix I):

Anger
Depression symptoms (collapsed)
Psychotic symptoms (collapsed)
Child abuse
Relationship dysfunction
Family support (subtracted)

Results for the total WRNA-T are also shown in Table 12. Coefficients are weaker for the WRNA-T than for the standalone WRNA because they are intended to be added to other, gender-neutral variables. Even so, correlations with serious and aggressive misconducts ranged from .16 to .32 in the construction validation, but shrank somewhat in the revalidation sample.

Predictive merits of the revised standalone scale are also seen in Table 13 for the Missouri and Ohio sites where it was tested. Correlations and AUC values were especially strong for the combined sample at the 12 month point, and for predictions of serious and aggressive misconducts. Additionally, AUC values for the Ohio sample were at or above .70 on all outcome measures, regardless of time frame. The Missouri results showed strong predictions of serious misconducts at 6 months, but results became much less stable at the 12 month point when the sample was considerably reduced in size.

Table 13 also showed a pattern suggesting that the gender responsive risk scale was more predictive, in most tests, than other models available to correctional agencies, including static custody scales and gender-neutral risk/needs scales.

This comparison is also seen in Table 14, where the results of augmenting the LSI-R with the WRNA-T are shown for the Rhode Island sample. The WRNA-T alone is significantly correlated with prison misconducts in 11 of 12 tests and makes a statistically significant contribution (partial correlation) to the LSI-R in 9 of 12 tests. It is also apparent that in a number of cases the overall predictive validity of the WRNA-T and the LSI-R is diminished by lower LSI-R correlations rather than the WRNA-T. Specifically in 8 of the 12 tests, correlations for the combined LSI-R/WRNA-T are lower than for the WRNA-T alone.

Finally, the WRNA-T alone is significantly correlated with prison misconducts for the sample as a whole and the Ohio sample in all tests and in 8 of 12 tests conducted on the Missouri sample.

Table 12. Bivariate Relationship between Revised Standalone WRNA and Misconducts, Construction and Revalidation Samples

	6 month (N=244/322)[a]						12 month (N=209/240)[b]					
	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
Scale	**Y/N**	**N**	**Y/N**	**N**	**Y/N**	**N**	**Y/N**	**N**	**Y/N**	**N**	**Y/N**	**N**
Validation Sample (N=244/209)												
Criminal History	.25***	.23***	.25***	.21***	.18***	.18***	.31***	.30***	.35***	.33**	.25***	.25***
AUC	.65		.68		.68		.68		.71		.69	
WRNA Stand Alone	.27***	.20***	.32***	.24***	.23***	.21***	.29***	.22***	.33***	.28***	.33***	.28***
AUC	.67		.74		.74		.68		.72		.77	
WRNA Stand Alone-Levels	.26***	.19***	.31***	.22***	.23***	.20***	.29***	..23***	.31***	.27***	.32***	.26***
AUC	.65		.71		.71		.62		.68		.73	
WRNA-T (Trailer)	.17**	.13***	.23***	.16***	.22***	.19***	.11**	.09*	.17***	.17***	.32***	.25***
Revalidation Sample (N=240/191)												
Criminal History							.17***	.14**	.24***	.21***	.17***	.12**
AUC							.61		.66		.65	
WRNA Stand Alone	.15***	.10*	.19***	.18***	.11**	--	.23***	.19***	.27***	.24***	.21***	.17***
AUC	.61		.66		.66		.65		.69		.69	
WRNA Stand Alone-Levels	.17***	.13***	.20***	.20***	.13**	.09*	.27***	.21***	.30***	.27***	.25***	.20***
AUC	.61		.65		.66		.66		.68		.70	
WRNA-T (Trailer)	.09*	.08*	.15***	.15***	.14***	.13***	.15**	.12**	.16***	.16***	.18***	.18***

***p≤.01
**p≤.05
*p≤.10
[c]Participants were in the 6 month window for at least 3 of the 6 months..
[b]Participants were in the 12 month window for at least 9 of the 12 months.

Table 13. Bivariate Relationship between Revised Cumulative Scales and Prison Misconducts, Ohio and Missouri.

	6 month						12 month					
	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N
Full Revised WRNA Assessment, (Missouri & Ohio (N=484/400)												
Static Custody	.16***	.14***	.16***	.15***	.11***	.09**	.25***	.23***	.29***	.28***	.21***	.19***
AUC	.61		.61		.62		.65		.69		.67	
Gender Neutral	.21***	.16***	.20***	.17***	.11***	.08**	.26***	.22***	.30***	.24***	.21***	.18***
AUC	.63		.65		.62		.65		.60		.68	
WRNA Stand Alone	.21***	.16***	.26***	.21***	.18***	.14***	.26***	.21***	.30***	.26***	.27***	.23***
AUC	.64		.70		.71		.67		.70		.73	
WRNA Stand Alone-Levels	.22***	.17***	.26***	.21***	.19***	.15***	.29***	.22***	.31***	.27***	.29***	.24***
AUC	.63		.68		.69		.66		.68		.71	
Missouri (N=98/53)												
Static Custody	--	--	--	--	--	--	--	.18*	.27**	.20*	--	--
AUC	.51		.51				.57		.64			
Gender Neutral	.21**	.13*	.21**	.15**	--	--	.20*	.18*	.25**	--	--	--
AUC	.62		.63				.63		.65			
WRNA Stand Alone	.25***	.18**	.33***	.26***	--	--	.18*	.22*	.18*	.19*	--	--
AUC	.65		.70				.62		.59			
WRNA Stand Alone-Levels	.20**	.18**	.28***	.22***	--	--	--	.20*	--	--	--	--
AUC	.62		.67		--		--		--		--	
Ohio (N=386/347)												
Static Custody	.21***	.18***	.19***	.18***	.13***	.11***	.26***	.27***	.29***	.29***	.24***	.23***
AUC	.65		.65		.63		.66		.69		.68	
Gender Neutral	.21***	.17***	.20***	.17***	.13***	.11***	.26***	.26***	.30***	.27***	.25***	.23***
AUC	.66		.61		.65		.66		.70		.70	
WRNA Stand Alone	.25***	.20***	.24***	.21***	.19***	.16***	.30***	.27***	.33***	.29***	.30***	.26***
AUC	.70		.71		.71		.70		.73		.74	
WRNA Stand Alone-Levels	.26***	.19***	.25***	.21***	.22***	.18***	.32***	.27***	.34***	.30***	.32***	.28***
AUC	.68		.69		.71		.69		.71		.73	

***p≤.01
**p≤.05
*p≤.10

Table 14. Bivariate Relationship between Revised Cumulative WRNA-Trailer Scales and Prison Misconducts.[a]

	6 month						12 month					
	Any Misc.		Ser. Misc.		Agg. Misc.		Any Misc.		Ser. Misc.		Agg. Misc.	
	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N	Y/N	N
WRNA Trailer, (Missouri, Ohio & Rhode Island (N=640/400)												
Cumulative WRNA-Trailer[a]	.13***	.11***	.18***	.15***	.19***	.17***	.12***	.11***	.17***	.17***	.26***	.22***
Missouri (N=98/53)												
Cumulative WRNA-Trailer[a]	.21**	.17**	.36***	.30***	.15*	.15*	--	.20*	--	.20*	--	--
Ohio (N=386/347)												
Cumulative WRNA-Trailer[a]	.23***	.18***	.23***	.19***	.19***	.16***	.23***	.19***	.25***	.21***	.24***	.21***
Rhode Island (N=154/69)												
LSI-R[b]	.15**	.21***	--	--	.15**	.15**	--	.28***	--	.25**	.32***	.34***
AUC	.54		.54		.74						.79	
Cumulative WRNA-Trailer[a]	.19**	.26**	.11*	.14**	.16**	.16**	--	.31***	.22**	.32**	.44***	.47***
LSI-R + WRNA-Trailer[c]	.17**	.25***			.17**	.17**	--	.31***	.17*	.29***	.38***	.40***
AUC	.54		.55		.77				.59		.80	
Partial WRNA-Trailer[d]	.13*	.18**	.12*	.12*			--	.19*	.17*	.22**	.33***	.36***

***p≤.01
**p≤.05
*p≤.10

[a] The WRNA Trailer is not intended to be a standalone risk assessment instrument, but rather a supplement to a gender neutral dynamic risk assessment, such as the LSI-R. Therefore, these correlations are not expected to be as high as a full risk/needs assessment.

[b] The LSI-R is a standalone risk assessment instrument correlations with outcomes at 12 months should be above .25.

[c] The LSI-R plus the WRNA trailer is a full risk/needs assessment. Correlations at 12 months should be above. 25. In some instances these are pulled down by low LSI-R values.

[d] This partial correlation shows the predictive validity (variation) that is attributable to the WRNA. It can be seen than in most instances, the WRNA makes a statistically significant contribution to that offered by the LSI-R, alone.

Risk/Need Scales in Prison Settings: Some Precautions

It is extremely important to alert potential users of the WRNA to the fact that many agencies using the WRNA in prison settings use it only as a needs assessment to aid the process of case management, treatment assignments and re-entry planning. Thus, the risk scales, shown in Tables 13 and 14, are seldom used to make assignments to minimum, medium, and maximum housing units even though our studies typically have found the WRNA scales to be more predictive of prison misconducts than static criminal history items which form the foundation of the current generation of custody classification tools. The WRNA is more predictive than criminal history items because it is comprised of measures such as mental health and trauma, and troubled inmates make difficult prison adjustments. For legal reasons, however, agencies typically chose against raising custody assignments on the basis of criteria associated with adversities that women have limited control over (see Wright, Van Voorhis, Salisbury & Bauman, 2012). Simply put, the observation that troubled women (high WRNA scores) make somewhat more difficult prison adjustments than women with high criminal history scores (high custody scores) is "an inconvenient truth."

A less common application of high WRNA scores involves agencies which view high WRNA scores as indicative of high need and then use such scores to assign women to treatment-intensive units of the facility. Of course, doing so also serves to address security issues because the treatment of risk factors is a form of risk management.[10] Moreover, a substantial number of women with high WRNA scores also have high custody scores, anyway.

[10] The majority of women who score high on prison custody classification also score high on the WRNA. The choice of which tool to use depends upon its intended purpose. Custody classification assessments cannot inform treatment programming decisions; they only address security considerations. Dynamic risk/needs assessments such as the WRNA and the LSI-R address both security and treatment considerations because they afford a risk scale *and* scores on individual needs.

Whether the Institutional WRNA risk score is used for housing assignments or not, it is highly recommended that community risk scores (derived from the Pre-release WRNA) be used to assist re-entry planning. Although the Pre-release WRNA involves the summation of different risk/need domains than the Institutional WRNA, re-entry case managers are advised not only to address important individual risk/needs but to do so with knowledge of which re-entering women offenders are at high risk of recidivism. These offenders should receive highest priority for community services, more intensive community supervision, and transitional housing.

CONCLUSION

Revalidation tests of the 2008 WRNA and WRNA-T, with no changes, produced acceptable results for 2 of 3 research sites, Ohio and Rhode Island. Results for the Missouri prison site were unacceptable. As noted, in the methodology section, however, it was difficult to secure the cooperation of prison interviewers and potential participants. Moreover, the sample size for the 12 month window was quite small (N=53) due to releases. This introduced a degree of instability to the findings. A report prepared specifically for Missouri officials encouraged officials to view prison results with caution (Van Voorhis et al., 2012).[11]

The present study tested a number of improvements to dynamic risk/need scales which resulted in improvements to the predictive validity of these specific risk/need scales as well as to the predictive validity of the assessment's institutional risk scale. Improvements were seen even for the Missouri sample, however, they were not as conclusive as those seen for Rhode Island or Ohio.

The study has succeeded in producing a somewhat shorter assessment than the original tool. This occurs primarily with the omission of two abuse survey scales measuring adult abuse

[11] Results for Missouri's prerelease and probation samples, however, were favorable.

(victimization) and child abuse. Most importantly, we have much more confidence in the stability of the assessment, because it now is seen to be predictive across several jurisdictions. Because, revisions were made using construction and revalidation samples, we have reduced the urgency for revalidation research. Nevertheless, revalidation tests by other researchers are strongly encouraged.

This study also afforded an opportunity to prepare a trailer (WRNA-T) for use with the LSI-R. In most tests, this tool significantly augmented the predictive validity of the LSI-R and provided a means for screening according to gender-responsive needs that are not contained on the LSI-R. A number of jurisdictions have chosen to use the WRNA-T solely as a needs assessment, thus avoiding the complication of adding the gender-responsive scales to the LSI-R and recalibrating risk levels. While that is a reasonable possibility, it was clear that the contribution of the WRNA-T to the predictive validity of the LSI-R was favorable (see Table 14).

We did not succeed in a goal of creating a single tool for use with pre-release, probation, and prison settings. However, the differences between the tools are not to the interviews or scales themselves but rather to the computation of total risk scores. The three tools each add a different subset of individual risk/need domains. For example, as noted above, we seldom see education, parental issues, and employment/financial issues impacting prison adjustment; these issues are more likely to appear on community risk scales because they are stressors that have a more direct effect on day to day functioning in the community. It is possible to accommodate these differences in risk scales through software computation formulas. The University of Cincinnati's Corrections Institute (UCCI) provides identical training protocols for each instrument, and the differences between interviews are minor.

Notwithstanding these contributions, there are precautions to be taken in understanding these findings. Follow-up time frames are limited to 12 months, both by the terms of NIC funding and by the releases of inmate participants. It is well known that longer follow-up time frames tend to produce better results, especially when they extend to a 24 month window. One reason for this is that short time frames achieve lower base rates of outcome occurrences. Low base rates typically attenuate findings. As shown in Table 3, this occurs primarily with the aggression outcome measures.

Though not shown in these analyses, results varied somewhat from interviewer to interviewer. Separate analyses found that some interviewers produced data which achieved lower predictive validity coefficients than others, especially on sensitive scales pertaining to abuse, trauma, and relationships. Further examination of these findings showed that these interviewers also incurred more missing data and were known by their colleagues to have been conducting their interviews too quickly. These are implications for both training protocols and staff selection.

Finally, we note that interviewers for the WRNA assessments were trained immediately prior to data collection. In contrast, the Rhode Island LSI-R interviewers had been trained several years prior to this study. State officials observed that many of the LSI-R interviewers were due to receive refresher training. The difference in proximity to training may explain why the WRNA-T was somewhat more predictive than the LSI-R. Dynamic assessments such as the WRNA and the LSI-R require careful monitoring for quality assurance; the validity of either assessment is likely to diminish when quality assurance becomes lax.

This report is the last of 6 reports prepared for NIC. Separate reports were prepared for Iowa, Rhode Island, and Missouri as their data became available. In addition, three reports

summarized findings for probation, prerelease, and prison settings. Each summary report produced a final instrument considered to be optimal for that setting. A total of 11 sites were involved. Only two of these, Ohio probation and Missouri prison samples, produced questionable findings. Both involved limited cooperation from staff in the respective agencies. A third, Iowa probation, provided a truncated distribution because probationers diagnosed as low risk were screened from participation in the Women Offender Case Management Project (WOCMM). As a result, the Iowa sample included very few low risk women, and findings were slightly more limited than those for other projects. The assessments were valid in all other sites as well as in the 3 sites studied during the 2008 construction validation research.

References

Andrews, D. A.& Bonta, J. (1995).*Level of Service Inventory-Revised.* North Tonawanda, NY: Multi-Health Systems.

Andrews, D. A.& Bonta, J. (2010).*The psychology of criminal conduct* (3rdEd.). Cincinnati, OH: Anderson.

Belknap, J. (2007). *The invisible woman: Gender, crime, and justice* (3rdEd.). Belmont, CA: Thompson Wadsworth.

Blanchette, K. (2004). Revisiting effective classification strategies for women offenders in Canada. *Feminism & Psychology, 14,* 231-236.

Blanchette, K. & Brown, S. L. (2006).*The assessment and treatment of women offenders: An integrative perspective.* Chichester: Wiley.

Bloom, B., Owen, B., & Covington, S. (2003). *Gender-responsive strategies: Research practice and guiding principles for women offenders*. Washington, DC: USDOJ, National Institute of Corrections.

Brennan, T.& Austin, J. (1997).*Women in jail: Classification issues.* Washington, DC: USDOJ, National Institute of Corrections.

Brennan, T. (1998). Institutional classification of females: Problems and some proposals for reform. In R. T. Zaplin (Ed.) *Female offenders: Critical perspectives and effective interventions* (pp. 179-204). Gaithersburg, MD: Aspen.

Brennan, T., Dieterich, W., & Oliver, W. (2006). *COMPAS: Technical manual and psychometric report Version 5.0.* Traverse City, MI: Northpointe Institute.

Browne, A., Miller, B., & Maguin, E. (1999). Prevalence and severity of lifetime physical and sexual victimization among incarcerated women. *International Journal of Law and Psychiatry, 22,* 301-322.

Chesney-Lind, M. (1997). *The female offender: Girls, women, and crime.* Thousand Oaks, CA: Sage.

Daly, K. (1992). Women's pathways to felony court: Feminist theories of lawbreaking and problems of representation. *Southern California Review of Law and Women's Studies, 2,* 11-52.

Farr, K. A. (2000). Classification for female inmates: Moving forward. *Crime and Delinquency, 46,* 3-17.

Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology, 34,* 575-607.

Hardyman, P.L., & Van Voorhis, P. (2004). *Developing Gender-Specific Classification Systems for Women Offenders*. Washington, DC: USDOJ, National Institute of Corrections.

Hewitt, J., Poole., E., & Regoli, R. (1984). Self-reported and observed rule-breaking in prison: A look at disciplinary response. *Justice Quarterly, 1(3)*: 437-447.

Maltz, M. (1984). *Recidivism.* Orlando, FL: Academic Press.

Morash, M. (2006). *Understanding gender, crime, and justice*. Thousand Oaks, CA: Sage.

Morash, M. (2010). *Women on Probation and Parole: A Feminist Critique of Community programs and Services.* Lebanon, NH: Northeastern University Press.

Morash, M., Bynum, T. S., & Koons, B. A. (1998). *Women offenders: Programming needs and promising approaches.* Washington, DC: USDOJ, National Institute of Justice.

Reisig, M. D., Holtfreter, K., & Morash, M. (2006). Assessing recidivism risk across female pathways to crime. *Justice Quarterly, 23*, 384-405.

Sherer, M., Maddus, J., Mercandante, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R. (1982). The Self Efficacy Scale: Construction and Validation. *Psychological Reports*, 51, 663-671.

Van Voorhis, P. (1994). Measuring Prison Disciplinary Problems: A Multiple Indicators Approach to Understanding Prison Adjustment. Justice Quarterly, 11(4): 679-709.

Van Voorhis, P. (2012). On behalf of women offenders: Women's place in the science of evidence-based practice. *Criminology and Public Policy*, 11:111-145.

Van Voorhis, P. & Presser, L. (2001).*Classification of women offenders: A national assessment of current practices.* Washington, DC: USDOJ, National Institute of Corrections.

Van Voorhis, P., Brushett, R. & Bauman, A., (2012). Revalidation of the Women's Risk/Needs Assessment: Results for the Missouri Department of Corrections, Probation, Prison, and Prerelease Samples. Cincinnati OH: Center for Criminal Justice Research.

Van Voorhis, P., Bauman, A., & Brushett, R. (2012). Revalidation of the Women's Supplemental Risk/Needs Assessment: Results for the Rhode Island Department of Corrections, Prison and Pre-release Samples. Cincinnati OH: Center for Criminal Justice Research.

Van Voorhis, P., Bauman, A., & Brushett, R. (2012). Revalidation of the Women's Risk/Needs Assessment: Pre-release Results. Cincinnati OH: Center for Criminal Justice Research.

Van Voorhis, P., Bauman, A., & Brushett, R. (2013). Revalidation of the Women's Risk/Needs Assessment: Probation Results. Cincinnati OH: Center for Criminal Justice Research.

Van Voorhis, P., Wright, E., Salisbury, E., & Bauman, A. (2010). Women's risk factors and their contributions to existing risk/needs assessment: The current status of a gender-responsive supplement. *Criminal Justice and Behavior*, *37*, 261-288.

Wright, E., Van Voorhis, P., Salisbury, E., and Bauman, A. (2012). Gender-Responsive Lessons Learned and Policy Implications for Women in Prison: A Review. *Criminal Justice and Behavior, 39,* 1612-1632.