Revalidation of the Women's Risk Needs Assessment:

Probation Results[1]

Final Report

January 2013

Patricia Van Voorhis, Ph.D.
University of Cincinnati

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

Rachel Brushett, M.S.
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 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:

Probation Results[2]

Final Report

January, 2013

Executive Summary

Patricia Van Voorhis, Ph.D.

University of Cincinnati

Ashley Bauman, M.S., M.B.A.

University of Cincinnati

Rachel Brushett, M.S.

University of Cincinnati

In order to facilitate appropriate treatment planning and risk management for women offenders, the National Institute of Corrections (NIC) and the University of Cincinnati (UC) engaged in a series of cooperative agreements that resulted in the development of two types of gender-responsive assessments. The work began in 1999 with a pilot study in the Colorado Department of Corrections and continued with three projects in Maui, Minnesota, and

[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 author and do not necessarily represent the official position or policies of the United States Department of Justice or the National Institute of Corrections.

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 most of the questions now contained on the trailer were 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 construction validation studies also produced different versions of the WRNA and the WRNA-T 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.

This report presents findings from a second cooperative agreement between the University of Cincinnati and the National Institute of Corrections. The present study was begun in 2009. 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 determine the level of shrinkage in predictive validity from the construction to revalidation studies. Additionally, the 2009 cooperative agreement sought to refine several of the dynamic risk/needs scales in order to further improve predictive validity. In doing so, this research tested a number of new items 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 construction validation research and afforded an opportunity to partition the

combined samples into construction and revalidation samples. The present report focuses on the Probation WRNA and WRNA-T. Two additional reports provide similar findings for female inmates in prison and prison re-entry settings.

The rationale for seeking to improve the Probation WRNA reflected the fact that the original tool was developed on small probation samples and had not been revalidated. In contrast, the present study secured a sample of 585 participants across across four sites, Missouri (N=91), Ohio (N=112), Iowa (N=329), and Minnesota (N=53). Even so, we had hoped for large random samples of participants, but instead we were forced to accept a number of sample irregularities. For example, the Iowa and Minnesota projects screened out lower risk offenders prior to administering the WRNA assessment. A more serious problem occurred in Ohio, where the study could not obtain sufficient cooperation from probation officers who were asked to recruit potential participants. The Missouri sample, which was considered to be a random sample, was small, because data collection began too late.

All sites furnished data for gender responsive scales. However, the gender-neutral scales could only be tested in Ohio and Missouri where the full-stand alone, WRNA was administered and tested (N=203). The following analytical steps were employed:

1. Individual risk/need scales developed during the original study were tested, through analysis of correlations (Pearson's r and AUC values) with outcome measures. These tests involved the same items and scoring protocols resulting from the 2004-2008 construction validation study. Analyses were run twice, once for the full assessment (WRNA) and once for the trailer assessment (WRNA-T).

2. The original total risk/needs score (developed through construction validation research), including risk levels, was tested on the research samples for the present study (Missouri and Ohio). Additionally, the cumulative WRNA-T scales were added to the LSI-R for the Iowa and Minnesota samples and tested for predictive validity.

3. 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 probation sites studied. The total sample was divided into a construction sample (N=101 for gender-neutral scales and 292 for

gender-responsive scales) and a revalidation sample (N=102 for gender-neutral scales and 293 for gender-responsive scales). The construction and revalidation samples were drawn through a systematic random selection process where every other case from the total pool of participants was selected for the construction sample, and the remaining cases were reserved for the revalidation sample. Items were developed on the construction validation sample and retested (confirmed) on the revalidation sample.

4. Because both the construction and revalidation samples were small, another analysis was conducted which tested the new scales for each state sample. Scales found to be predictive in the construction validation study but not in the revalidation study, could nevertheless be retained for the final assessment if they were found to be predictive in two or more of the state samples. This rather unusual procedure accommodates some degree of sample specificity which occurred for both the WRNA and the LSI-R risk/need scales. That is, there was a tendency for a given risk/need domain to be predictive in some samples and not in others, regardless of the assessment used (LSI-R or WRNA).

5. Selection of a final risk/needs stand-alone and trailer scales considered both the results for the construction and revalidation samples (step 3) and the state- specific findings (step 4). The WRNA and the WRNA-T were developed in the construction validation sample, retested on the revalidation sample, and then tested for specific sites.

Offense-Related Outcome Measures

Most participants were followed up for 12 months, and results were reported at a 6 month interval and for the entire 12 months. Because probationers could fail (recidivate) in a variety of ways, a number of outcome measures were examined: a) NEW ARRESTS (Y/N); 2) NEW CONVICTIONS (Y/N); 3) INCARCERATIONS (through technical or law violations or new arrests/convictions); 4) TECHNICAL VIOLATIONS (Y/N); 5) ANY OFFENSE-RELATED FAILURE (e.g., new arrests/convictions as well as behavior which could have been processed as a violation but through officer discretion or agency policy was nevertheless recorded as a violation), and 6) ANY FAILURE (any of the above). As can be seen from Table 1, there was considerable variation across sites as to how offenders failed. In Missouri, for example, offenders were far more likely to receive violations than arrests and convictions. There are very few arrests and even fewer convictions in the Missouri follow-

up data, and incarcerations were often through a violation process. As such, measures of arrests and convictions are not as meaningful for that sample as they were for Ohio, Iowa, or Minnesota. By the 12 month follow-up, some Minnesota follow-up measures appeared to be redundant. Finally, the INCARCERATION measure was only available for the Missouri sample. The Ohio data were obtained from county web sites and may be incomplete. The inconsistency across sites renders state-specific findings very important. There were no measures that could be considered comparable across sites. Additionally, the present study considered only 12 months of follow-up. It is likely that the preferred 24 months of follow-up would have further improved base rates and the prospects for even stronger findings.

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

Site	Arrests		Conv.		Ret Pris[a]		Tech. Viol.		Offense Fail		Any Fail	
	N	%	N	%	N	%	N	%	N	%	N	%
6 Month Follow-up[b]												
Missouri (N=91)	5	5.5	2	2.2	10	11.1	55	60.4	20	22.0	59	64.8
Ohio (N=112)	12	10.7	7	6.3			6	5.4	15	13.4	18	16.1
Iowa (N=329)	20	6.1	13	4.0			45	13.7	28	8.5	53	16.1
Minnesota (N=53)	19	35.8	7	13.2			16	30.2	20	37.7	21	39.6
12 Month Follow-up[c]												
Missouri (N=85)	9	10.6	2	2.4	16	18.8	63	74.1	29	34.1	66	77.6
Ohio (N=102)	20	19.6	15	14.7			11	10.8	24	23.5	27	26.5
Iowa (N=316)	48	15.2	30	9.5			87	27.5	60	19.0	95	30.1
Minnesota (N=51)	22	43.1	13	25.5			23	45.1	23	45.1	25	49.0

[a]Data on incarcerations were only available for the Missouri sample.
[b] At least 3 months of follow-up was required to be included in the 6 month follow-up window.
[c] At least 8 months of follow-up was required to be included in the 12 month follow-up window.

Results

Revalidation of Individual Risk/Need Scales Created During the 2008 Construction Validation Research

Findings for the combined samples (Tables 2 and 3) appeared to confirm the scales developed during the 2008 construction validation research. For the combined samples, all but 1 (mental health history) of 22 gender-responsive measures were related to at least one outcome measure at probability levels of .05 or lower during the 6 month follow up period. Three measures failed to correlate at this level during the 12 month follow-up (depression,

psychosis, and sexual abuse). Findings for the mental health scales generally were weak, but emerged in some of the state-specific findings. The strongest predictors among the gender-responsive risk/needs scales included economic issues (employment/financial) and anger. Importantly, cognitions pertinent to anger and self-efficacy were more potent predictors of offense-related outcomes than more widely used cognitive measures of antisocial thinking. This was true regardless of whether or not antisocial thinking was measured by the LSI-R or the WRNA. Abuse risk/needs scales and measures of housing safety were predictive in many instances, but to a somewhat lessor degree than other measures.

Strengths were also worthy of note. Self-efficacy, parental involvement, and educational assets were negatively associated with new offenses and violations. As such they proved to be sources of resilience for these participants. Relationship support and family support showed weaker results.

The gender-neutral risk/needs scales generally did not correlate with outcomes as strongly as the gender-responsive risk/needs scales regardless of whether they were measured by the LSI-R or the WRNA. The exceptions were consistent with other WRNA research and involved substance abuse and education and employment scales.

State-specific findings (shown in Tables 7 through 10 of the report) were generally stronger for Missouri and Minnesota than for Ohio and Iowa. Findings for the Ohio sample were especially weak, implicating the sample selection process as well as the validity of the follow-up measures.

Table 2. Bivariate Relationship between LSI-R and Original WRNA Assessment Scales and 6 Month Recidivism, All Jurisdictions.[a]

Scale	Arrests Y/N	Conv. Y/N	Incarc[b] Y/N	Tech. V. Y/N	Offense Fail Y/N	Any Fail Y/N
LSI-R (N=382)						
Criminal history	.07*	.13***		.09*	.12***	.13***
Attitudes						
Education/employment	.17***	.11***		.14***	.20***	.18***
Financial						
Accommodations						
Leisure/recreation						
Family/marital	.08*				.07*	
Antisocial friends	.09**	.08*		.11***	.10**	.14***
Alcohol/ drugs	.07*			.13***	.08*	.10**
Emotional/personal				.08*		
WRNA Scales-Gender Neutral (N=203)[c]						
Criminal history	-.15*			.14**		
Attitudes	.17**			-.22***		-.13**
Educational needs				.24***		.20***
Antisocial friends			.16*			
Substance abuse history			.18**	.27****		.23***
Substance abuse (current)			.43***	.37***	.13**	.30***
WRNA-Gender Responsive (N=585)[d]						
Educational assets (strength)	-.11***			-.14***	-.12***	-.16***
Employment/financial	.14***		.23***	.18***	.14***	.19***
Housing safety	.12***	.08*	.46***	.12***	.16***	.11***
Anger	.21***	.07*		.09**	.14***	.10*
Mental health history		.08*		.06*	.07*	
Depression (symptoms)		.06*	.25***		.07**	.07**
Psychosis (symptoms)			.22**			
Child abuse	.14***	.08*		.06*	.12***	.09**
Adult abuse	.11**	.09**		.07**	.09***	.08**
Sex abuse (adult or child)	.11***				.10***	.07**
Physical abuse (adult or child)	.15***	.13**		.09**	.12***	.10***
PTSD		.05*	.26***	.07*	.08**	.06*
Parental difficulties (Interview)		-.07*	.19**			
Parental stress (all)	.06*		.24***	.08**	.07*	.09**
Parental involvement[e]	-.12**	-.12***	-.29***	-.20***	-.17***	-.17***
Family conflict	.13***	.08**			.10***	.06*
Family support (strength)	-.07**	-.08**	-.21**		-.08*	
Relation. satisfaction (strength)		-.05*		-.06*	-.06*	-.07**
Relationship dysfunction			.25***			
Child abuse (survey)	.09**	.10**			.07**	
Adult victimization (survey)	.11***	.07*			.09**	.10***
Self-efficacy (strength)	-.09**	-.07*	-.22**	-.16***	-.09**	-.15***

[a]To be included in the 6 month follow-up frame, participants needed to be at least 3 months post-interview.
[b] Revoked to prison data was available only for the Missouri probation sample (N=91).
[c] The sample size reduced to 112 (the Ohio cases) on the arrest and convictions measures, with the exclusion of Missouri cases. Missouri arrests and convictions evidenced extremely low base rates.
[d] The sample size reduced to 494 on the arrest and convictions measures, with the exclusion of Missouri, data arrests and convictions evidenced extremely low base rates.
[e] The scales pertains to mothers of children under 18 (N=357). On the revoked to prison measure, N=59.
***p≤.01, **p≤.05 , *p≤.10

Table 3. Bivariate Relationship between LSI-R and Original WRNA Assessment Scales and 12-Month Recidivism, All Jurisdictions.[a]

Scale	Arrests Y/N	Conv. Y/N	Incarc.[b] Y/N	Tech.Viol. Y/N	Offense Fail Y/N	Any Fail Y/N
LSI-R (N=366)						
Criminal history	.12***	.11**		.17***	.15***	.17***
Attitudes						
Education/employment	.23***	.14***		.21***	.25***	.21***
Financial		.07*			.07*	
Accommodations		.10**		.12***		.12***
Leisure/recreation						
Family/marital	.09**				.08*	
Antisocial friends		.08*		.14***		.13***
Alcohol/ drugs	.11**			.17***	.10**	.15***
Emotional/personal	.08*			.08*		
WRNA Scales-Gender Neutral (N=187)[c]						
Criminal history			.14*	.10*		
Attitudes	.14*			-.27***		-.14**
Educational needs		-.14*		.23***		.21***
Antisocial friends			.22**			
Substance abuse history			.21**	.25***	.12*	.22***
Substance abuse (current)	.14*		.39***	.34***	.22***	.33***
WRNA-Gender Responsive (N=553)[d]						
Educational assets (strength)	-.12***			-.16***	-.12***	-.17***
Employment/financial	.13***	.07*	.26***	.20***	.15***	.21***
Housing safety	.08**	.06*	.37***	.12***	.11***	.09**
Anger	.25***	.11***		.16***	.21***	.17***
Mental health history	.07*			.08*	.09**	.06*
Depression (symptoms)				.06*		.05*
Psychosis (symptoms)				.06*		
Child abuse	.12***	.08**			.11***	
Adult abuse	.07*	.06**		.08**	.07**	.08**
Sex abuse (adult or child)					.06*	
Physical abuse (adult or child)	.15***	.11***		.10***	.13***	.10***
PTSD			.20**			
Parental difficulties (Interview)		-.07*	.20**			
Parental stress (all)			.21**	.08*		.08**
Parental involvement[e]	-.22***	-.18***	-.32***	-.20***	-.23***	-.20***
Family conflict	.11***	.07*			.06*	
Family support (strength)	-.08**	-.09**	-.16*		-.09***	
Relation. satisfaction (strength)				-.13***	-.09**	-13***
Relationship dysfunction	.06*		.22**	.06*		
Child abuse (survey)	.12***	.09**			.08**	
Adult victimization (survey)	.10***	.12***		.09*	.09**	.10***
Self-efficacy (strength)	-.12***	-.08*	-.31***	-.18***	-.14***	-.16***

[a]To be included in the 12 month follow-up frame, participants needed to be at least 8 months post-interview.
[b] Revoked to prison data was available only for the Missouri probation sample (N=85).
c The sample size reduced to 102 (the Ohio cases) on the arrest and convictions measures with the exclusion of Missouri cases. Missouri arrests and convictions evidenced extremely low base rates.
[d] The sample size reduced to 468 on the arrest and convictions measures, with the exclusion of Missouri data.
[e] The scales pertains to mothers of children under 18 (N=334). On the revoked to prison measure, N=53.
***p≤.01, **p≤.05, *p≤.1

Revalidation of the Original Cumulative Risk Scales Created During Construction Validation Research (2004-2008).

Construction validation research, completed in 2008, developed risk scores for the stand alone WRNA by summing the following risk/need scales: criminal history, antisocial attitudes, employment/financial, housing safety, anger, antisocial friends, psychotic symptoms, depression/anxiety, substance abuse history, current substance abuse, family conflict, and parental stress (see Appendix B). Strengths pertaining to educational assets, family support, and self-efficacy were subtracted from the total. It was possible to retest these scales in Ohio and Missouri. Results are shown in Table 4, below.

The findings were favorable for the Missouri site but not for the Ohio site. For the Missouri sample, the WRNA scale was predictive of returns to prison and any failure at both the 6 and 12 month follow-up time periods. Results for arrests and convictions are not shown due to limited variation on the outcome variables. It was not surprising that the results were not acceptable for the Ohio sample. As noted above, both the sampling process and the collection of follow-up data were compromised.

Results for the trailer sites, Iowa and Minnesota, are shown in Table 5. With the exception of the conviction measure, the WRNA-T showed high correlations with 12 month outcomes in the combined samples. Partial correlations were also significant, indicating that the WRNA_T offered significant incremental validity to the prediction offered by the LSI-R. Bivariate correlations between the WRNA-T alone and outcomes were stronger for the Minnesota sample than the Iowa sample. Failure to find significant incremental validity on 3 of 15 tests is likely to be attributable to the low base rates on the conviction outcome measure. Trailer results may also be seen in Table 4, because it was possible to extract just the trailer scales from the stand alone WRNA. Results were favorable for Missouri but not Ohio.

Table 4. Predictive Validity of Stand-alone WRNA, Original 2008 Construction Validation Scales (Missouri and Ohio)

Scale	Arrests[a]		Conv.[a]		Incar.[b]		Tech. Viol.		Offense Fail		Any Fail	
	Y/N	AUC	Y/N	AUC	Y/N	AUC	Y/N	AUC	Y/N	AUC	Y/N	AUC
Revalidation Original Instrument – 6 months[c]												
WRNA Stand-alone-All (N=203)							.25***	.65	.12**	.58	.23***	.63
Levels-All (N=203)							.23***	.63	.15**	.60	.24***	.63
WRNA-T- All (N=203)							.16***		.14**		.18**	
WRNA Stand-alone -Missouri (N=91)					.40***	.83	.34***	.69	.19**	.60	.30***	.67
Levels-Missouri (N=91)					.35***	.81	.31***	.67	.15*	.60	.27***	.65
WRNA-T- Missouri (N=91)					.34***		.30***		.16*		.27***	
WRNA Stand-alone -Ohio (N=112)	--		--				--	.64	--		--	
Levels-Ohio (N=112)	.13*						--	.64	.12*	.60	.18**	.64
WRNA-T-Ohio (N-112)	--		.13*				--					
Revalidation Original Instrument - 12 Months[c]												
WRNA Stand-alone- All(N=187)							.24***	.63	.17***	.60	.24***	.63
Levels (N=187)							.24**	.63	.19***	.61	.25***	.63
WRNA-T-All (N=187)							.21***		.14***		.20***	
WRNA Stand-alone -Missouri (N=85)					.37***	.75	.30***	.68	.15*	.57	.27***	.67
Levels-Missouri (N=85)					.32***	.72	.31***	.69	.15*	.59	.28***	.68
WRNA-T-Missouri (N=85)					.27***		.27***		--		.27***	
WRNA Stand-alone -Ohio (N=102)	--		--				.16**	.64	.17**	.60	.18**	.60
Levels-Ohio (N=102)	.17**	.62	--				.16*	.65	.20**	.62	.20**	.62
WRNA-T-Ohio (N=102)	--		--				.21***		.16**		.16 *	

***p≤.01; **p≤.05; *p≤.10

[a] Arrest and conviction data are not tested for Missouri due to limited variation.

[b] Incarcerations are only tested for Missouri.

[c] Participants required at least 3 month of follow-up to be included in the 6 month data and at least 8 months to be included in the 12 month da

Table 5. Predictive Validity of LSI-R and the Original WRNA-T, 2008 Construction Validation Scales (Minnesota and Iowa)

Scale	Arrests		Conv.		Incar.[a]		Tech. Viol.		Offense Fail		Any Fail	
	Y/N	AUC	Y/N	AUC	Y/N	AUC	Y/N	AUC	Y/N	AUC	Y/N	AUC
Revalidation Original Instrument – 6 months[b]												
Minnesota and Iowa (N=382)												
LSI-R (N=382)	.14***	.65	.12***	.66			.16***	.63	.17***	.66	.18***	.63
WRNA-T	.25***		.09*				.19***		.24***		.21***	
LSI-R + WRNA-T	.22***	.72	.13***	.68			.20***	.66	.23***	.71	.22***	.67
Partial corr.	.19***		.07*				--		.18***		.16***	
Minnesota (N=53)												
LSI-R	.23**	.65	--				.28**	.69	.22*	.63	.22**	.63
WRNA-T	.25***		--				.21**		.21**		--	
LSI-R + WRNA-T	.26**	.67	--				.27***	.69	.23**	.67	.21**	.63
Partial corr.	--		--				--		--		--	
Iowa (N=329)												
LSI-R	.14***	.69	.13***	.71			.14***	.62	.19***	.71	.18***	.64
WRNA-T	.14***		--				.14***		.15***		.16***	
LSI-R + WRNA-T	.16***	.72	.11**	.68			.16***	.63	.20***	.72	.20***	.66
Partial corr.	.08*		--				.08*		.08*		.10**	
Revalidation Original Instrument - 12 Months[b]												
Minnesota and Iowa (N=367)												
LSI-R	.21***	.68	.13***	.62			.23***	.66	.22***	.66	.22***	.64
WRNA-T	.26***		.10**				.24***		.25***		.24***	
LSI-R + WRNA-T	.27***	.70	.13***	.62			.27***	.66	.27***	.68	.26***	.67
Partial corr.	.18***		--				.15***		.17***		.16***	
Minnesota (N=51)												
LSI-R	.28***	.67	--				.33***	.68	.27**	.66	.26**	.64
WRNA-T	.36***		--				.31***		.33***		.27***	
LSI-R + WRNA-T	.34***	.68	--				.35***	.69	.32***	.66	.29**	.64
Partial corr.	.24**		--				--		.20**		--	
Iowa (N=316)												
LSI-R	.22***	.69	.14***	.65			.22***	.66	.22***	.68	.22***	.65
WRNA-T	.16***		--				.19***		.18***		.20***	
LSI-R + WRNA-T	.23***	.70	.13***	.63			.25***	.66	.24***	.68	.25***	.68
Partial corr.	.08*		--				.11**		.09*		.12**	

***p≤.01; **p≤.05 ;*p≤.10

[a] Incarcerations are only tested for Missouri.

[b] Participants required at least 3 month of follow-up to be included in the 6 month data and at least 8 months to be included in the 12 month data.

Revision of the WRNA Scales

Improvement of the assessment scales was achieved through the development and testing of new items on a construction validation sample and then revalidating those items. Presumably, the revalidation of the new scales will reduce concerns for the need of another revalidation study. The new items and item analyses resulted in improvements to 11 of 29 need domains: Criminal History, Antisocial Friends, Substance Abuse History, Current Substance Abuse, Employment/Financial, Depression (symptoms), Family Conflict, Parental Involvement, Parental Stress, Relationship Support, and Relationship Dysfunction.

Compilation of revised WRNA cumulative scales followed a process similar to that used for the individual scales. Scales were first developed for a construction validation sample and then retested on a revalidation sample. The optimal scales for the revised WRNA stand-alone and WRNA-T are below:

WRNA Stand-Alone	**WRNA-T**
Criminal History	*Depression (collapsed)*
Antisocial Friends	*Employment/Financial*
Substance Abuse History	*Housing Safety*
Current Substance Abuse	*Anger*
Depression (collapsed)	*Child Abuse*
Employment/Financial	*Adult Abuse*
Housing Safety	*Parental Stress (collapsed)*
Anger	*Educational Assets (subtracted)*
Child Abuse	*Self-Efficacy (collapsed and subtracted)*
Adult Abuse	*Family Support (collapsed and subtracted)*
Parental Stress (collapsed)	
Educational Assets (subtracted)	
Self-Efficacy (collapsed and subtracted)	
Family Support (collapsed and subtracted)	

Validation Results for the WRNA-T

Results for the construction and revalidation samples are shown in Table 6, below. Scale refinements represented an improvement to the predictive validity of the trailer in comparison to results shown in Tables 4 and 5 above for the 2008 WRNA-T. The cumulative scale also showed stability over the construction and revalidation samples. Moreover, the revised trailer cumulative scale contains items pertaining to abuse (child and adult), which did not appear in the original trailer. The revised scale also does not include a risk factor identifying psychotic behaviors. This was seldom predictive and is included in Part IV of the new assessment.

An additional examination apart from the construction and revalidation samples is shown in Table 7. Here results are examined for the LSI-R/trailer sites (Iowa and Minnesota). Table 7 shows some improvement in the predictive validity of the revised WRNA-T over the original (results which are shown in Table 5). However, the improvement was largely attributable to the Minnesota results. Results for Iowa were remarkably similar to those shown for the 2008 trailer. In addition, the overall findings for Iowa were less favorable than those for Minnesota. We note in the report, however, that the Iowa participants were receiving intensive treatment for many of the risk factors assessed by the WRNA-T. Treatment accomplishments on any of the risk factors could have attenuated these results.

Table 6. Bivariate Relationship between Revised WRNA Trailer (WRNA-T) and Outcomes for Construction and Revalidation Samples, All Sites.

	Construction (N=276)						Revalidation (N=277)					
Scale	Arrests[a] Y/N	Conv.[a] Y/N	Incar.[b] Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N	Arrests[a] Y/N	Conv[a]. Y/N	Incar. [b] Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N
6 Month Follow-up (N=292, N=293)												
WRNA_T	.25***	.10*	--	.18***	.22***	.23***	.27***	.13**	.52***	.26***	.27***	.26***
12 Month Follow-up (N=276, N=277)												
WRNA-T	.25***	.11**	--	.24***	.24***	.25***	.26***	.17***	.46***	.28***	.26***	.26***

[a] Analysis of arrest and conviction data excluded Missouri cases, because of limited variation. . Analysis at 6 months included 248 in the construction sample and 246 in the revalidation sample. Analysis at 12 month included 236 in the construction sample and 232 in the revalidation sample.

[b] Analysis is only for Missouri cases. Analysis at 6 months included 44 in the construction sample and 47 in the revalidation sample. Analysis at 12 month included 40 in the construction sample and 45 in the revalidation sample.

***p≤.01; **p≤.05 ;*p≤.10

Table 7. Bivariate Relationship between LSI-R and Revised WRNA Trailer (WRNA-T and Outcomes for Iowa and Minnesota.

	Arrests		Conv.		Incar.[b]	Tech. Viol.		Offense Fail		Any Fail	
Scale	**Y/N**	**AUC**	**Y/N**	**AUC**	**Y/N**	**Y/N**	**AUC**	**Y/N**	**AUC**	**Y/N**	**AUC**
6 Month Follow-up											
Total (N=382)											
LSI-R	.14***	.65	.12***	.66		.16***	.63	.17***	.66	.18***	.63
WRNA-T	.30***		.12***			.22***		.29***		.25***	
LSI-R+WRNA-T	.24***	.73	.14**	.69		.22***	.67	.26***	.73	.24***	.68
Partial corr.	.26***		.07*			.17***		.24***		.19***	
Iowa (N=329)											
LSI-R	.14***	.69	.13***	.71		.14***	.62	.19***	.71	.18***	.64
WRNA-T	.16***		--			.14***		.18***		.18***	
LSI-R+WRNA-T	.17***	.72	.12**	.70		.16***	.64	.21***	.73	.21***	.67
Partial corr.	.10**		--			.09*		.10**		.11**	
Minnesota (N=53)											
LSI-R	.23**	.65	--	.60		.28***	.69	.22*	.63	.22*	.63
WRNA-T	.36***		--			.34***		.32***		.28***	
LSI-R+WRNA-T	.30***	.70	--	.61		.33***	.73	.28**	.68	.26**	.67
Partial corr.	.28***		--			.21*		.23**		.18*	
12 Month Follow-up											
Total (N=366)											
LSI-R	.21***	.68	.13***	.62		.23***	.66	.22***	.66	.22***	.64
WRNA-T	.29***		.16***			.26***		.29***		.26***	
LSI-R+WRNA-T	.28***	.71	.16***	.65		.28***	.67	.29***	.70	.28***	.67
Partial corr.	.22***		.11***			.18***		.22***		.19***	
Iowa (N=315)											
LSI-R	.22***	.70	.14***	.65		.22***	.66	.22***	.68	.22***	.65
WRNA-T	.18***		.10**			.21***		.20***		.21***	
LSI-R+WRNA-T	.23***	.70	.14***	.65		.25***	.66	.25***	.69	.25***	.66
Partial corr.	.09**		--			.12**		.11**		.12***	
Minnesota (N=51)											
LSI-R	.28**	.67	--	.58		.33***	.69	.27**	.66	.26**	.64
WRNA-T	.44***		--			.37***		.40***		.36***	
LSI-R+WRNA-T	.37***	.71	--	.59		.37***	.71	.35***	.69	.32***	.67
Partial corr.	.36***		--			.21*		.31***		.25**	

***p≤.01; **p≤.05 ;*p≤.10

Validation Results for the Stand-Alone WRNA

Results for the tests of the WRNA stand-alone assessment on construction and revalidation samples are shown in Table 8. Predictive validity for the total scale was strong for measures of incarceration (Missouri), technical violations, and any failure, at the 12 month follow-up point. Results found in the construction sample were confirmed in the revalidation sample.

State-specific findings appear in Table 9. Results generally showed an improvement over the 2008 assessments. Nevertheless, these findings varied by state and outcome measure. Predictions of returns to prison, technical violations, offense-related failures, and any type of failure were strong for the Missouri sample. Correlations were high, and AUC values equaled or approach .70. However, results for the Ohio sample were unacceptable and indicative of errors in the sample selection process and/or the collection of follow-up data.

Conclusion

Validation tests of both the 2008 WRNA and WRNA-T as well as the revised instruments produced acceptable results that afford a good deal of confidence in these tools. 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. The Ohio findings stand in contrast to the generally favorable findings for the other three sites, however. As noted earlier, the study, which was voluntary, was terminated for lack of participation on the part of probation officers and potential participants.

The larger 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

Table 8. Bivariate Relationship between Revised WRNA Stand-Alone and Outcomes for Construction and Revalidation Samples, All Sites.

	Construction						Revalidation					
Scale	Arrests[a] Y/N	Conv.[a] Y/N	Incar.[b] Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N	Arrests Y/N	Conv. Y/N	Incar.[b] Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N
6 Month Follow-up (N=101, N=102)												
WRNA-Stand Alone	--	--	.34**	.31***	.14*	.32***	--	--	.60***	.40***	.18**	.29***
12 Month Follow-up (N=94, N=93)												
WRNA-Stand Alone	.26***	--	.39***	.35***	.27***	.42***	--	--	.57***	.32***	.16*	.19**

[a] Analysis of arrest and conviction data excluded Missouri cases, because of limited variation. Analysis at 6 months included 57 in the construction sample and 55 in the revalidation sample. Analysis at 12 month included 54 in the construction sample and 48 in the revalidation sample.

[b] Analysis is only for Missouri cases. Analysis at 6 months included 44 in the construction sample and 47 in the revalidation sample. Analysis at 12 month included 40 in the construction sample and 45 in the revalidation sample.

***p≤.01; **p≤.05; *p≤.10

Table 9. Bivariate Relationship between the WRNA Stand-Alone Assessment and Outcomes For Missouri and Ohio.

	Arrests		Conv.		Incar.[b]		Tech.Viol.		Offense Fail		Any Fail	
Scale	**Y/N**	**AUC**	**Y/N**	**AUC**	**Y/N**	**AUC**	**Y/N**	**AUC**	**Y/N**	**AUC**	**Y/N**	**AUC**
6 Month Follow-up												
Total (N=203)												
WRNA							.35***	.71	.16***	.59	.31***	.67
Levels							.36***	.70	.13**	.58	.29***	.65
WRNA-T							.20***		.15**		.22***	
Missouri (N=91)												
WRNA					.48***	.89	.37***	.71	.26***	.64	.33***	.68
Levels					.41***	.82	.38***	.70	.24***	.64	.33***	.68
WRNA-T					.34***		.30***		.16***		.29***	
Ohio (N=112)												
WNRA	--		--				.13*	.67	--	--	--	--
Levels	--		--				.12*	.67	--	--	--	--
WRNA-T	--		--				--		--	--	--	--
12 Month Follow-up												
Total (N=187)												
WRNA							.34***	.69	.22***	.62	.31***	.67
Levels							.35***	.69	.22***	.62	.31***	.66
WRNA-T							.24***		.17***		.24***	
Missouri (N=85)												
WRNA					.46***	.80	.32***	.70	.24**	.62	.28***	.67
Levels					.39***	.76	.36***	.69	.25***	.64	.29***	.68
WRNA-T					.28***		.28***		--		.27***	
Ohio (N=102)												
WRNA	--	--	--	--			.17**	.65	.16**	.60	.20**	.62
Levels	--	--	--	--			.16**	.63	.14*	.58	.16**	.59
WRNA-T							.17**		.18**		.18**	

***p≤.01; **p≤.05; *p≤.10

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 avoided 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 validity of the LSI-R as a prediction was favorable (see Tables 6 & 7).

Notwithstanding these contributions, there are some necessary precautions to be taken in understanding these findings. The more ideal research sample would have involved a random statewide sample, or several of them. Two of the three sites sampled for this study, truncated assessment distributions, through a process which attempted to screen-out low risk women. A third probation site (Ohio) was affected by poor cooperation from probation officers in referring women to the study and by concerns for the validity of the follow-up measures. The one site (Missouri) which tapped all potential, English-speaking clients was delayed in starting, a fact which reduced the size of the sample for that site.

Follow-up data are also likely to be truncated. The follow-up time period for the present study was 12 rather than 24 months. The earlier 2004 - 2008 studies found more impressive results at 24 months than at 12. Limited base rates are known to attenuate findings, and longer follow-up periods improve base rates, which in turn tends to improve predictive validity coefficients. Of greatest concern in this regard involves mental health scales. In other studies, these often did not appear to emerge as correlates until the 18 to 24 month time frame.

With the exception of Iowa, where the assessment was used for case planning for many women offenders, the study samples are rather small. The present study amassed data on 203 cases for gender-neutral variables and 585 cases for gender-responsive variables. This necessitated a boot-strap approach where scales were developed in a construction validation sample and retested in a revalidation sample as well as in state-specific samples.

Though not shown in these analyses, we also detected some interviewer effects. 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 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, in some tests results for the LSI-R trailer were not as favorable as those for the WRNA stand-alone instrument. We note that interviewers for the WRNA assessments were trained immediately prior to data collection. In contrast a number of state officials observed that many of the LSI-R interviewers were due to be retrained. 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.

Even with these limitations, results are somewhat more favorable than typically seen at a 12 month follow-up. It is likely that the study limitations did not bias findings in a favorable direction. Typically, limitations with base rates, sample size, and quality assurance attenuate findings. Most importantly, we have much more confidence in the stability of the assessment, because it now is seen to be predictive across several jurisdictions.

Revalidation of the Women's Risk Needs Assessment:

Probation Results[3]

Final Report

January 2013

Patricia Van Voorhis, Ph.D.
University of Cincinnati

Rachel Brushett, M.S.
University of Cincinnati

Ashley Bauman, M.S., M.B.A.
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

Clearly, the present report is not solely the work of its three authors. Behind the scenes were funders, leaders of correctional agencies, interviewers, trainers, UC research staff, and research, planning, and agency staff at the research sites. All worked hard to bring this research to fruition.

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 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); Michael Walton, Esq., Court Administrator at Hamilton County Probation; Robert Veatch, ISP Supervisor at Hamilton County Probation; Julie Frey, Director of Probation at Clermont County Common Pleas Court; Julie Rud, Policy, Planning, and Evaluation Manager (MN DOC), Lettie Prell, Director of Research (IA DOC), and Anne Brown, Key Coordinator of IA WOCMM. 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 was depended upon the assistance of staff in research and planning departments at three of the sites, Iowa, Missouri, and Minnesota. We worked extensively with Julie Rud at MNDOC, Lettie Prell at IADOC, and with David Oldfield and Fred Martin at MDOC. 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, Emily Wright, 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, Noreen Loftus-Spilman, and Lindsay Morrow.

We are especially appreciative to the 500 plus 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 lend 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.

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, 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 were 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 serving 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) which stressed 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 women's pathways perspective asserted that women's unique needs were not adequately tapped by the current generation of risk/needs assessments, instruments 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.

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. Additionally, the 2009 cooperative agreement sought to refine several of the WRNA 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 construction

validation research and afforded an opportunity to partition the combined samples into construction and revalidation samples.

The present report focuses on the Probation WRNA, the assessment intended for use at intake and follow-up case-management of women under community supervision. The present study secured a sample of 585participants across four sites, Missouri, Ohio, Iowa, and Minnesota. In sum, the specific goals of the present study were as follows:

1. To revalidate the 2008 individual risk/needs scales as well as the total risk scale resulting from the sum of all predictive risk/needs scales.

2. To test the contributions of new test items to the predictive validity of specific risk/needs scales as well as to the total risk scale representing the sum of risk/need factors predictive of offense-related outcomes.

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

4. 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 to 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 across settings.

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

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. Following focus groups with correctional personnel and female offenders, the Missouri Women's Issues Committee refined items and questions into the draft used in the 2008 construction validation study. The MDOC used an in-house static risk calculation to assess both their male and female populations, so the addition of the WRNA as a risk/needs assessment for their female population was a marked change in the way that female offenders and their unique needs were approached. The MDOC intended the standalone WRNA to serve two functions:

1. As a means of identifying high need female probationers for assignment to treatment-intensive units/caseloads while under supervision; and
2. To facilitate case planning by identifying risk/need factors that needed to be addressed in order to promote success and reduce future offending.

At the beginning of 2008, Missouri was poised to begin statewide implementation of the assessment. However, as implementation efforts were about to begin, a number of factors contributed to significant delays. Changes in leadership temporarily halted the process. A decision to await automation of the assessment delayed the process further while software was developed. Budget cuts also played a role and delayed the training of staff. As set-backs persisted, staff commitment decreased. Moreover, they viewed a 45 minute assessment to be a significant increase in workload compared to their current static risk calculation (which provided no treatment-relevant guidance). Finally, it was determined that rather than proceed with full implementation, Missouri would engage in a pilot project. This

would give select staff a chance to use the assessment and evaluate its utility. Leadership hoped this pilot period would help staff re-invest in the project.[4]

Piloting of the WRNA began in the spring of 2010. Prior to the start of the pilot period, the University of Cincinnati Corrections Institute conducted the standard WRNA training on evidence-based practice, gender-responsive principles and practice, administration and scoring of the assessment, interview skills, and assessment-based case management. Select probation officers in 18 of the 43 probation districts participated in this training program and were selected to pilot the assessment. They were instructed to use the assessment with female offenders on their caseloads immediately following training. They would then report back to leadership on their experiences with the assessment allowing leadership and University of Cincinnati research staff to address their concerns.

Their assessments were used in the current revalidation study. With original plans for full state implementation no longer proceeding as planned and a limited time to complete data collection per the study funding deadlines, a much smaller number of assessments were collected for the research study than initially anticipated. For this study, assessment data were collected over 5 months, netting a total of 91 cases.[5] De-identified data was transmitted to University researchers allowing for a 100% response rate.

[4] The pilot proved worthwhile. MDOC implemented the WRNA statewide in November of 2012.
[5]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.

Ohio

Data for this research-only sample were collected at two probation departments in Southwestern, Ohio, from April 5 through July 13, 2010.[6] The standalone version of the WRNA was tested. The first site, Hamilton County Probation, served the Greater Cincinnati Area. The probation department agreed to refer female clients to the study following their standard probation visit. The clients were informed of the nature of the study and asked to participate. Researchers quickly encountered three problems with this approach. First, the number of female clients estimated by the probation department to be under probation supervision was much larger than the number of women actually reporting to their probation officers. Second, not all of the probation officers referred their clients to the researchers and others were doing so on an inconsistent basis. Third, unlike the sites where the WRNA was the required intake assessment, participation at this site was voluntary. Eligible probationers had competing demands (i.e., work, family, transportation) and were often not able to participate in the research due to these other commitments. Finally, data problems were compounded by the lack of access to offender files needed to corroborate current and prior offense data. After one month of interviewing, researchers determined that their initial goal of obtaining 400 interviews at this site was not going to be possible during the study time frame. They sought additional cases from a neighboring county, Clermont County, Ohio, but similar problems occurred. As a result, work at these sites was discontinued in the fall of 2010. The resulting sample of 112 cases (102 from Hamilton County and 10 from Clermont

[6]IRB approval was granted for this study (#10022302) in March 2010. Re-approvals have been granted annually.

County) is likely to be non-representative of all women serving probation in these two counties.

As will be seen, research staff were also concerned with the web-based follow-up data available at these two sites. Web-based systems are maintained by county courts and it is unclear if these are updated on a consistent basis. Failure to update the systems in a timely manner could have resulted in some offenses not being captured by the researchers. The report proceeds with an analysis of these data. However, results must be viewed with caution.

Iowa

Implementation of the Women's Risk Needs Assessment - Trailer (WRNA-T) began in January 2009, and interviews for this study were administered between March 24, 2009 and April 19, 2011. Women probationers in four Iowa probation districts were interviewed using the LSI-R (Andrews &Bonta, 1995) and the WRNA-T as part of the districts' Women Offender Case Management (WOCMM)(Van Dieten, 2008) project. The WOCMM caseloads were gender-responsive caseloads designed to provide a gender-responsive, holistic approach to case management (see Van Dieten, 2008 or www.nicic.org for more information). Women were intended to be screened into the WOCMM program by scoring moderate to high risk on either the Iowa Risk Assessment or the Level of Service Inventory – Revised.[7] Those women then received the WRNA-T probation/parole trailer to inform case planning. All WRNAs were conducted by WOCMM probation officers. The screening requirements appeared to reduce the number of low-risk women admitted into the study.

[7] At the time, the LSI-R was administered to both male and female probationers throughout the Iowa probation districts.

Although risk scores ranged from 7 to 46, only 10.3 percent of the participants were classified as low risk or low/moderate risk on the LSI-R.

Not all of the women who were assessed went on to participate in the WOCMM program; however, all assessments were included in the study, netting a total of 329 assessments.[8] Of the 329 participants, 256 (77.8 percent) were WOCMM participants and 73 (22.2 percent) were not.

Prior to implementing the WRNA, experienced probation officers received extensive training in the WOCMM as well as the standard WRNA training. All participants, regardless of WOCMM participation, were interviewed by the WOCMM officers.

Minnesota

Like Missouri, Minnesota was one of the research sites under the previous NIC cooperative agreement. In April 2009, two gender-responsive probation officers in Hennepin County, Minnesota adopted the WRNA-T to supplement the LSI-R which was used with their caseloads at that time. One of these officers was trained as an agency trainer in the WRNA by the University of Cincinnati in January 2010. Both officers administered the WRNA-T to women offenders on their caseloads who scored 24 or higher on the LSI-R. The LSI-R and the WRNA-T were also administered to select women supervised by a special unit designed to target misdemeanor probation offenders. In this unit, two probation officers supervised female offenders who had 4 or fewer prior prostitution convictions on their record

[8] IRB approval for the Iowa study (#10122704) was granted in February 2011. Re-approval was granted in 2012.

and were charged by the City of Minneapolis for a new prostitution offense. The results were then used with both groups to facilitate probation case planning and service provision for women offenders. For this study, all assessments completed from April 10, 2009 through July 31, 2010 were included, netting a total of 53 cases. De-identified data was transmitted to University researchers allowing for a 100% response rate.[9]

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

In sum, due to many of the constraints imposed by research in action settings, the samples cannot be considered representative of all women serving probation terms in these sites. Participants in Iowa and Minnesota were admitted to their respective samples after a process which screened out lowest risk offenders. In both cases, however, the samples still evidenced some low risk participants. The Ohio sample was adversely impacted by limited cooperation from probation officers who were asked to refer women to the study and by some concern for the consistency of the follow-up data. The Missouri sample was not affected by any of these considerations, but was nevertheless quite small. Findings, while favorable, must be viewed with these considerations.

[9]The Minnesota study was reviewed and approved (#10081001) by the Institutional Review Board (IRB) at the University of Cincinnati in August 2010. Re-approvals have been granted annually.

METHODOLOGY

A total of 585 women offenders participated in the probation study. Of those, 203 completed the full, stand-alone WRNA (91 in Missouri and 112 in Ohio) and 382 completed the LSI-R and the WRNA-T (329 in Iowa and 53 in Minnesota).

Sample Description

Table 1 presents demographic and criminal history characteristics for the two samples that were utilized for the examination of the WRNA stand-alone assessment (Missouri and Ohio). As shown in Table 1, the samples were comparable in terms of age, marital and parental status, and employment. The average ages of women in the two samples ranged from 32.6 years of age (Ohio) to 34.2 years of age (Missouri). In both of the samples, less than a quarter of the participants were married at the time of the study (20.2 percent in Missouri and 19.8 percent in Ohio). Much larger percentages of each group had minor children, including 64.8 percent of the Missouri sample and 62.2 percent of the Ohio sample.

The participants in the Ohio sample were more likely to have a high school diploma (68.8 percent) and were more likely to be employed (51.8 percent) than the participants in the Missouri sample (54.9 percent and 47.3 percent).

Table 1. Frequency and Percent Distribution of Demographic and Offense-Related Background Characteristics of Samples Participating in the Validation of the Stand-Alone WRNA.

	Missouri		Ohio	
Characteristic	**N**	**Percent**	**N**	**Percent**
	91	100.0	112	100.0
Age	N = 84			
18-20 years old	5	6.0	8	7.6
21-30 years old	26	31.0	45	42.9
31-40 years old	32	38.1	27	25.7
41-50 years old	16	19.0	19	18.1
51 years and older	5	6.0	6	5.7
	$\overline{X}$ = 34.2 yrs		$\overline{X}$ = 32.6 yrs	
Race	N = 89		N = 111	
Asian	0	0.0	0	0.0
African American	30	33.0	57	50.9
Hispanic/Latina	0	0.0	0	0.0
Native American	1	1.1	1	0.9
Other	0	0.0	0	0.0
White	58	63.7	53	47.3
Currently Married	N = 84		N = 111	
Yes	17	20.2	22	19.8
Client Have Children Under 18			N = 111	
Yes	59	64.8	69	62.2
Employment				
Employed (full or part-time, child care, student, or disabled)	43	47.3	58	51.8
Not employed	48	52.7	54	48.2
H.S. Grad or GED				
Yes	50	54.9	77	68.8

Table Continues

Table 1. Frequency and Percent Distribution of Demographic and Offense-Related Background Characteristics of Samples Participating in the Validation of the Stand-Alone WRNA, continued

	Missouri		Ohio	
Characteristic	**N**	**Percent**	**N**	**Percent**
	91	100.0	112	100.0
Most Serious Present Offense	N = 89			
Arson	0	0.0	4	3.6
Assault	0	0.0	4	3.6
Burglary	8	9.0	9	8.0
Damage property	0	0.0	0	0.0
Dangerous drugs	36	40.4	31	27.7
DWI	2	2.2	0	0.0
Family offenses	5	5.6	5	4.5
Forgery/Fraud	15	16.9	12	10.7
Homicide/Manslaughter	1	1.1	0	0.0
Kidnapping	0	0.0	0	0.0
Larceny	15	16.9	35	31.3
Other	3	3.4	4	3.6
Robbery	0	0.0	3	2.7
Sex offenses	0	0.0	1	0.9
Stolen property	3	3.4	2	1.8
Traffic offenses	0	0.0	0	0.0
Weapon offenses	1	1.1	2	1.8
Present Offense Violent	N = 89			
Yes	2	2.2	16	14.3
Prior Felonies			N = 109	
Yes	29	31.9	27	24.8
Prior Incarcerations				
Yes	36	39.6	11	9.8

Racial distributions differed across samples with the majority of the Ohio sample being African American (50.9 percent) and the majority of the Missouri sample being white (63.7 percent).The Ohio sample admitted more participants with a current violent offense, 14.3 percent versus 2.2 in Missouri. In both samples, drug-related offenses, larceny, and forgery/fraud were the most common current offense; however, participants in Missouri were more likely to be convicted of a drug offense (40.4 percent) while participants in Ohio were more likely to be convicted of larceny (31.3 percent). Women in Missouri had more extensive criminal backgrounds than women in the Ohio sample. This held true for both prior felonies (31.9 percent in Missouri and 24.8 percent in Ohio) and prior incarcerations (39.6 percent in Missouri and 9.8 percent in Ohio). However, the differences in prior records could be attributable to the lack of access to official records in Ohio. In this regard, the proportion of Ohio participants with prior prison terms appears to be somewhat suspect.

Table 2 presents demographic and criminal history characteristics for the two LSI-r and WRNA-T samples (Iowa and Minnesota). As can be seen, in Table 2, the samples differed much more than the Missouri and Ohio samples. Age distributions were similar (34.0 years in Iowa and 33.3 years in Minnesota), but considerable differences emerged on other demographic and offense history measures.

The largest racial groups in the Iowa sample were white (79.3 percent) and African American (19.5 percent). In the Minnesota sample African American probationers comprised 60.4 percent of the sample and Native Americans 34.0 percent.

Women in the Iowa sample were more likely to be married than women in the Minnesota sample (30.8 percent compared to 14.0 percent); however, the women in

Minnesota were more likely to have children younger than 18 (70.6 percent compared to 59.6 percent in Iowa). Over three-quarters of the women in Iowa (77.1 percent) possessed a high school diploma or general equivalency diploma compared to only 62.3 percent of women in the Minnesota sample. Similarly, women in Iowa were more likely to be employed at the time of the study (50.0 percent compared to 32.1 percent in Minnesota).

Women in Iowa were most likely to be convicted of a drug-related offense (48.6 percent) while women in Minnesota were most likely to be convicted of a sex-related (prostitution) offense (58.5 percent). Violent offenses were uncommon in both samples (10.9 percent in Iowa and 13.2 percent in Minnesota). Again, difference existed when looking at prior criminality. Women in Minnesota were more likely to have had a prior felony on their record (41.5 percent) than women in Iowa (5.5 percent), but women in Iowa were more likely to have had a prior incarceration (62.9 percent) than those in Minnesota (3.8 percent). Differences may reflect a misunderstanding of whether or not to count prior jail time as well as a failure or inability to verify self-report responses to the question regarding prior incarcerations. The revised WRNA and revisions to the training protocol attempt to correct this error.

Differences between the two samples were likely due to the assessment eligibility requirements for the different agencies. As noted above, some of the Minnesota participants were assessed due to offense-related criteria while Iowa participants were assessed based on LSI-R scores.

Table 2. Frequency and Percent Distribution of Demographic and Offense-Related Background Characteristics of Samples Participating in the Validation of the Supplemental (Trailer) WRNA.

	Iowa		Minnesota	
Characteristic	**N**	**Percent**	**N**	**Percent**
	329	100.0	53	100.0
Age			N = 52	
18-20 years old	27	8.2	2	3.8
21-30 years old	103	31.3	20	38.5
31-40 years old	113	34.3	17	32.7
41-50 years old	71	21.6	10	19.2
51 years and older	15	4.6	3	5.8
	$\overline{X}$ = 34.0 yrs		$\overline{X}$ = 33.3 yrs	
Race				
Asian	1	0.3	0	0.0
African American	64	19.5	32	60.4
Hispanic/Latina	0	0.0	3	5.7
Native American	3	0.9	18	34.0
Other	0	0.0	0	0.0
White	261	79.3	0	0.0
Currently Married	N = 318		N = 50	
Yes	98	30.8	7	14.0
Client Have Children Under 18	N = 324		N = 51	
Yes	193	59.6	36	70.6
Employment	N = 324			
Employed (full or part-time, child care, student, or disabled)	162	50.0	17	32.1
Not employed	162	50.0	36	67.9
H.S. Grad or GED	N = 327			
Yes	252	77.1	33	62.3

Table Continues

Table 2. Frequency and Percent Distribution of Demographic and Offense-Related Background Characteristics of Samples Participating in the Validation of the Supplemental (Trailer) WRNA, continued

	Iowa		Minnesota	
Characteristic	**N**	**Percent**	**N**	**Percent**
	329	100.0	53	100.0
Most Serious Present Offense				
Arson	0	0.0	0	0.0
Assault	21	6.4	6	11.3
Burglary	12	3.6	2	3.8
Damage property	2	0.6	1	1.9
Dangerous drugs	160	48.6	5	9.4
DWI	27	8.2	0	0.0
Family offenses	0	0.0	0	0.0
Forgery/Fraud	41	12.5	1	1.9
Homicide/Manslaughter	1	0.3	1	1.9
Kidnapping	0	0.0	0	0.0
Larceny	39	11.9	4	7.5
Other	21	6.4	2	3.8
Robbery	0	0.0	0	0.0
Sex offenses	1	0.3	31	58.5
Stolen property	0	0.0	0	0.0
Traffic offenses	4	1.2	0	0.0
Weapon offenses	0	0.0	0	0.0
Present Offense Violent				
Yes	36	10.9	7	13.2
Prior Felonies				
Yes	18	5.5	22	41.5
Prior Incarcerations	N = 328			
Yes	207	62.9	2	3.8

Analytic Process

As noted above, the goals of the present study involved validating of the original Probation WRNA and WRNA-T, and examining ways to improve separate risk/need domains as well as 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 new offenses committed up to 12 months following each participants interview. Two such scales were examined, a stand-alone WRNA and a WRNA-T. The WRNA-T was designed as a supplement to gender-neutral risk assessments, such as the Northpointe COMPAS or the LSI-R, with gender-responsive scales. In the present study, the WRNA-T was tested as trailer for the LSI-R.[10]

The following analytical steps were employed:

1. Individual risk/need scales developed for the 2004-2008 construction validation samples were tested, through analysis of correlations and AUC values with outcome measures. These tests involved the same items and scoring protocols resulting from the 2004-2008 construction validation study. Analyses were run twice, once for the full assessment and once for the WRNA-T assessment. See Appendix A for the 2008 WRNA, Appendix B for the scoring form for the WRNA; Appendix C for the WRNA-T, and Appendix D for the scoring form of the original WRNA-T.

2. The original total risk/needs score (developed through construction validation research), including risk levels, was tested on the stand-alone WRNA samples for the present study (Missouri and Ohio). Additionally, the cumulative WRNA-T scales were added to the LSI-R for the Iowa and Minnesota samples and tested for predictive validity. For the WRNA-T sites, incremental validity was also tested. Incremental validity refers to the issue of whether the gender-responsive scales make a statistically significant contribution to the validity of the LSI-R.

3. 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

[10] In order to avoid use of redundant scales, the composition of the WRNA Trailer was specific to the gender-neutral assessment being used. For example, the Northpointe COMPAS did not contain mental health scales. Therefore, the COMPAS Trailer 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.

tested on a split half sample of all probation sites studied. The total sample was divided into a construction sample (N=101 for gender-neutral scales and 292for gender-responsive scales) and a revalidation sample (N=102 for gender-neutral scales and 293 for gender-responsive scales). The construction and revalidation samples were drawn through a systematic random selection process where a pool of all cases was available and every other case was selected for the construction sample, and the remaining cases reserved for the revalidation sample. Items 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 3. Table 3 shows very similar distributions across samples, and no significant differences on any of the background data tested.

4. Because both the construction and revalidation samples were small, another analysis was conducted which tested the new scales for each state sample. Scales found to be predictive in the construction validation study but not in the revalidation study, could nevertheless be retained for the final assessment if they were found to be predictive in two or more of the state samples. This rather unusual procedure accommodates some degree of sample specificity which occurred for both the WRNA and the LSI-R risk/need scales. That is, there was a tendency for a given risk/need domain to be predictive in some samples and not in others, regardless of the assessment used (LSI-R or WRNA). Therefore, we endeavored to create a total assessment that would be predictive across samples even if the most predictive need domains varied from sample to sample.

5. Selection of a final risk/needs stand-alone and WRNA-T scales considered both the results for the construction and revalidation samples (step 3) and the state- specific findings (step 4). Total risk/needs scales were developed in the construction validation sample, retested on the revalidation sample, and then tested for specific sites.

Data analysis employed bivariate correlations (Pearson's r) and analysis of receiver operating characteristics (AUC). 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. Missing data were replaced at scale medians and unless specifically noted did not surpass 5 percent of the cases.

Table 3. Frequency and Percent Distribution of Demographic and Offense-Related Background Characteristics of the Construction and Revalidation Samples.

Characteristic	Construction		Revalidation		Total	
	N	Percent	N	Percent	N	Percent
Age	N=248		N=246		N=494	
18-20 years old	13	5.2	24	9.8	37	7.5
21-30 years old	93	37.5	77	31.3	170	34.4
31-40 years old	76	30.6	83	33.7	159	32.2
41-50 years old	56	22.6	45	18.3	101	20.4
51 years and older	10	4.0	17	6.9	27	5.5
	$\bar{X}$ = 34.5 yrs		$\bar{X}$ = 33.4 yrs		$\bar{X}$ = 33.9 yrs	
Race	N = 291		N=292		N=583	
Caucasian	185	63.6	188	64.4	373	64.0
African American	90	30.9	93	31.8	183	31.4
Hispanic/Latina	1	0.3	2	0.7	3	0.5
Asian American	1	0.3	0	0.0	1	0.2
Native American	14	4.8	9	3.1	23	3.9
Currently Married	N = 282		N=281		N=563	
Yes	65	23.0	79	28.1	144	25.6
Client Has Children Under 18	N=288		N=289		N=577	
Yes	181	62.8	175	60.9	356	61.9
Employment	N=288		N=292		N=580	
Employed full time	54	18.8	49	16.8	103	17.8
Part-time or unable	89	30.9	88	30.1	177	30.5
Not employed	149	50.3	155	53.1	300	51.7
H.S. Grad or GED	N=292		N=293		N=585	
Yes	203	69.5	211	72.0	414	70.8
Most Serious Present Offense	N=292		N=292			
Violent	33	11.3	28	9.6	61	10.4
Property	107	36.6	99	33.9	206	35.3
Drug	109	37.7	121	41.4	230	39.4
Public order	38	13.0	33	11.3	71	12.2
Other	5	1.7	11	3.8	16	2.7
Prior Felonies	N = 102		N=102		N=204	
Yes	34	33.3.0	30	29.4	64	31.4

Some consideration was given to whether or not the Ohio data should be removed from further analysis, especially in tests where the Ohio data were included with other samples. The construction and revalidation process was of particular concern. When these analyses were run a second time without the Ohio data, however, similar findings were observed. As a result, the choice was made to include the Ohio data.

Offense-Related Outcome Measures

Most participants were followed-up for 12 months, and results were reported at a 6 and 12 month intervals. Because probationers could fail (recidivate) in a variety of ways, a number of outcome measures were examined: a) NEW ARRESTS (Y/N); 2) NEW CONVICTIONS (Y/N); 3) INCARCERATIONS (through technical or law violations or new arrests/convictions); 4) TECHNICAL VIOLATIONS; 5) ANY OFFENSE-RELATED FAILURE (e.g., new arrests/convictions as well as behavior which could have been processed as a violation but through officer discretion or agency policy was nevertheless recorded as a violation), and 6) ANY FAILURE (any of the above).

As can be seen from Table 4, there was considerable variation across sites as to how offenders failed. In Missouri, for example, offenders were far more likely to receive violations than arrests and convictions. There are very few arrests and even fewer convictions in the Missouri follow-up data, and incarcerations were often through a violation process. As such, measures of arrests and convictions are not as meaningful for that sample as they were for Ohio, Iowa, or Minnesota. By the 12 month follow-up, some Minnesota follow-up measures appeared to be redundant. Finally, the INCARCERATION measure was only available for the Missouri sample.

The methods for transmitting follow-up data also varied by site. Missouri and Iowa department research personnel downloaded electronic data from the agency's information system. Minnesota's follow-up data were collected by the interviewers and transmitted via an Access database. Ohio follow-up data collection involved UC research personnel examining county court websites, where offenses were recorded and made publically available. Officials reported that arrests and convictions were likely to be more accurate than the records for technical violations. The low number of technical violations reported for the Ohio sample confirms this observation.

The inconsistency across sites renders state-specific findings very important. There were no measures that could be considered comparable across sites. Additionally, the present study considers only 12 months of follow-up. It is likely that the preferred 24 months of follow-up would have further improved base rates and the prospects for even stronger findings.

Table 4. Follow-up Measures by Time Frame and Site

Site	Arrests		Conv.		Ret Pris[a]		Tech. Viol.		Offense Fail		Any Fail	
	N	%	N	%	N	%	N	%	N	%	N	%
6 Month Follow-up[b]												
Missouri (N=91)	5	5.5	2	2.2	10	11.1	55	60.4	20	22.0	59	64.8
Ohio (N=112)	12	10.7	7	6.3			6	5.4	15	13.4	18	16.1
Iowa (N=329)	20	6.1	13	4.0			45	13.7	28	8.5	53	16.1
Minnesota (N=53)	19	35.8	7	13.2			16	30.2	20	37.7	21	39.6
12 Month Follow-up[c]												
Missouri (N=85)	9	10.6	2	2.4	16	18.8	63	74.1	29	34.1	66	77.6
Ohio (N=102)	20	19.6	15	14.7			11	10.8	24	23.5	27	26.5
Iowa (N=316)	48	15.2	30	9.5			87	27.5	60	19.0	95	30.1
Minnesota (N=51)	22	43.1	13	25.5			23	45.1	23	45.1	25	49.0

[a]Data on incarcerations were only available for the Missouri sample.
[b] At least 3 months of follow-up was required to be included in the 6 month follow-up window.
[c] At least 8 months of follow-up was required to be included in the 12 month follow-up window.

RESULTS

Revalidation of Risk/Need Scales Created During 2008 Construction Validation Research

Correlations between the original 2008 risk/needs scales and various outcome measures are shown in Table 5 for 6 month outcomes and in Table 6 for 12 month outcomes. Results for the LSI-R scales are also shown in the top 10 rows of findings. State-specific findings appear in Tables 7 and 8 for the WRNA stand-alone sites (Missouri and Ohio) and in Tables 9 and 10 for the LSI-R sites (Minnesota and Iowa). As noted in the tables, arrest and conviction data were not examined for Missouri,[11] and incarcerations were only available for Missouri. Assessments tested for this portion of the analyses appear in Appendices A through D.

Findings for Combined Samples

Findings for the combined samples (Tables 5 and 6) appear to confirm the scales developed during the 2008 construction validation research. For the combined samples, all but 1 (mental health history) of 22 gender-responsive measures were related to at least one outcome measure at probability levels of .05 or lower during the 6 month follow up period. Three measures failed to correlate at this level during the 12 month follow-up (depression, psychosis, and sexual abuse).Findings for the mental health scales generally were weak, but emerged in some of the state-specific findings. The strongest predictors among the gender-responsive risk/needs scales included economic issues (employment/financial) and anger.Importantly, cognitions pertinent to anger and self-efficacy were more potent predictors of offense-related outcomes than more widely used cognitive measures of antisocial thinking. This was true regardless of whether or not antisocial thinking was measured by the LSI-R or the WRNA. Abuse risk/needs scales and measures of housing safety were predictive in many instances, but to a somewhat lessor degree than other measures.

[11]The Missouri probation districts appeared to process potential arrests and convictions through revocations, violations and incarceration. Base rates for arrests and convictions were extremely low.

Table 5. Bivariate Relationship between LSI-R and Original WRNA Assessment Scales and 6-Month Recidivism, All Jurisdictions.[a]

Scale	Arrests Y/N	Conv. Y/N	Incarc[b] Y/N	Tech. V. Y/N	Offense Fail Y/N	Any Fail Y/N
LSI-R (N=382)						
Criminal history	.07*	.13***		.09*	.12***	.13***
Attitudes						
Education/employment	.17***	.11***		.14***	.20***	.18***
Financial						
Accommodations						
Leisure/recreation						
Family/marital	.08*				.07*	
Antisocial friends	.09**	.08*		.11***	.10**	.14***
Alcohol/ drugs	.07*			.13***	.08*	.10**
Emotional/personal				.08*		
WRNA Scales-Gender Neutral (N=203)[c]						
Criminal history	-.15*			.14**		
Attitudes	.17**			-.22***		-.13**
Educational needs				.24***		.20***
Antisocial friends			.16*			
Substance abuse history			.18**	.27****		.23***
Substance abuse (current)			.43***	.37***	.13**	.30***
WRNA-Gender Responsive (N=585)[d]						
Educational assets (strength)	-.11***			-.14***	-.12***	-.16***
Employment/financial	.14***		.23***	.18***	.14***	.19***
Housing safety	.12***	.08*	.46***	.12***	.16***	.11***
Anger	.21***	.07*		.09**	.14***	.10*
Mental health history		.08*		.06*	.07*	
Depression (symptoms)		.06*	.25***		.07**	.07**
Psychosis (symptoms)			.22**			
Child abuse	.14***	.08*		.06*	.12***	.09**
Adult abuse	.11**	.09**		.07**	.09***	.08**
Sex abuse (adult or child)	.11***				.10***	.07**
Physical abuse (adult or child)	.15***	.13**		.09**	.12***	.10***
PTSD		.05*	.26***	.07*	.08**	.06*
Parental difficulties (Interview)		-.07*	.19**			
Parental stress (all)	.06*		.24***	.08**	.07*	.09**
Parental involvement[e]	-.12**	-.12***	-.29***	-.20***	-.17***	-.17***
Family conflict	.13***	.08**			.10***	.06*
Family support (strength)	-.07**	-.08**	-.21**		-.08*	
Relation. satisfaction (strength)		-.05*		-.06*	-.06*	-.07**
Relationship dysfunction			.25***			
Child abuse (survey)	.09**	.10**			.07**	
Adult victimization (survey)	.11***	.07*			.09**	.10***
Self-efficacy (strength)	-.09**	-.07*	-.22**	-.16***	-.09**	-.15***

[a]To be included in the 6 month follow-up frame, participants needed to be at least 3 months post-interview.
[b] Revoked to prison data was available only for the Missouri probation sample (N=91).
[c] The sample size reduced to 112 (the Ohio cases) on the arrest and convictions measures, with the exclusion of Missouri cases. Missouri arrests and convictions evidenced extremely low base rates.
[d] The sample size reduced to 494 on the arrest and convictions measures, with the exclusion of Missouri, data arrests and convictions evidenced extremely low base rates.
[e] The scales pertains to mothers of children under 18 (N=357). On the revoked to prison measure, N=59.
***p≤.01
**p≤.05
*p≤.10

Table 6. Bivariate Relationship between LSI-R and Original WRNA Assessment Scales and 12-Month Recidivism, All Jurisdictions.[a]

Scale	Arrests Y/N	Conv. Y/N	Incarc.[b] Y/N	Tech.Viol. Y/N	Offense Fail Y/N	Any Fail Y/N
LSI-R (N=366)						
Criminal history	.12***	.11**		.17***	.15***	.17***
Attitudes						
Education/employment	.23***	.14***		.21***	.25***	.21***
Financial		.07*			.07*	
Accommodations		.10**		.12***		.12***
Leisure/recreation						
Family/marital	.09**				.08*	
Antisocial friends		.08*		.14***		.13***
Alcohol/ drugs	.11**			.17***	.10**	.15***
Emotional/personal	.08*			.08*		
WRNA Scales-Gender Neutral (N=187)[c]						
Criminal history			.14*	.10*		
Attitudes	.14*			-.27***		-.14**
Educational needs		-.14*		.23***		.21***
Antisocial friends			.22**			
Substance abuse history			.21**	.25***	.12*	.22***
Substance abuse (current)	.14*		.39***	.34***	.22***	.33***
WRNA-Gender Responsive (N=553)[d]						
Educational assets (strength)	-.12***			-.16***	-.12***	-.17***
Employment/financial	.13***	.07*	.26***	.20***	.15***	.21***
Housing safety	.08**	.06*	.37***	.12***	.11***	.09**
Anger	.25***	.11***		.16***	.21***	.17***
Mental health history	.07*			.08*	.09**	.06*
Depression (symptoms)				.06*		.05*
Psychosis (symptoms)				.06*		
Child abuse	.12***	.08**			.11***	
Adult abuse	.07*	.06**		.08**	.07**	.08**
Sex abuse (adult or child)					.06*	
Physical abuse (adult or child)	.15***	.11***		.10***	.13***	.10***
PTSD			.20**			
Parental difficulties (Interview)		-.07*	.20**			
Parental stress (all)			.21**	.08*		.08**
Parental involvement[e]	-.22***	-.18***	-.32***	-.20***	-.23***	-.20***
Family conflict	.11***	.07*			.06*	
Family support (strength)	-.08**	-.09**	-.16*		-.09***	
Relation. satisfaction (strength)				-.13***	-.09**	-13***
Relationship dysfunction	.06*		.22**	.06*		
Child abuse (survey)	.12***	.09**			.08**	
Adult victimization (survey)	.10***	.12***		.09*	.09**	.10***
Self-efficacy (strength)	-.12***	-.08*	-.31***	-.18***	-.14***	-.16***

[a]To be included in the 12 month follow-up frame, participants needed to be at least 8 months post-interview.
[b] Revoked to prison data was available only for the Missouri probation sample (N=85).
c The sample size reduced to 102 (the Ohio cases) on the arrest and convictions measures with the exclusion of Missouri cases. Missouri arrests and convictions evidenced extremely low base rates.
[d] The sample size reduced to 468 on the arrest and convictions measures, with the exclusion of Missouri data.
[e] The scales pertains to mothers of children under 18 (N=334). On the revoked to prison measure, N=53.
***p≤.01
**p≤.05
*p≤.10

Strengths were also worthy of note. Self-efficacy, parental involvement, and educational assets were negatively associated with new offenses and violations. As such they proved to be sources of resilience for these participants. Relationship support and family support showed much weaker results.

The gender-neutral risk/needs scales generally did not correlate with outcomes as strongly as the gender-responsive risk/needs scales regardless of whether they were measured by the LSI-R or the WRNA. The exceptions were consistent with other WRNA research and involved substance abuse and education and employment scales.

State-Specific Findings

The Missouri and Ohio sites tested only the WRNA risk/need scales. The LSI-R was not administered. Tables 7 and 8 underscore concerns for the Ohio sample. Voluntary and limited participation along with potentially incomplete web-based outcome measures were likely to have attenuated the findings. Results were much stronger in states where the WRNA was actually in day to day use for risk assessment and treatment planning. By 12 months the findings for Ohio showed only a minimal improvement. The findings for Ohio are best described as attenuated. Many were just short of reaching significance, but insignificant just the same.

Results for states testing the LSI-R and using only the WRNA gender-responsive risk/needs scales are shown in Tables 9 and 10. These further underscore the fact that Ohio results were an exception to the other 3 sites. Across all sites, however, it can be seen that very few measures predicted across all samples, whether they were obtained through the WRNA or the LSI-R.

The patterns shown in Tables 9 and 10 are similar to those seen for the stand alone WRNA. The most important risk factors appeared to involve financial and educational issues as well as anger. Abuse measures were also strongly associated with recidivism in the two trailer sites as was substance abuse, when measured by the WRNA scale. Strengths pertaining to parental involvement and self-efficacy were also important sources of resilience.

Table 7. Bivariate Relationship between Original WRNA Assessment Scales and 6-Month Recidivism, Missouri and Ohio.[a]

	Missouri: WRNA (N=91)						Ohio: WRNA (N=112)					
	Arrests[b]	Conv.[b]	Incarc.	Tech. Viol.	Offense Fail	Any Fail	Arrests	Conv.	Incarc.	Tech. Viol.	Offense Fail	Any Fail
Scale	**Y/N**	**Y/N**	**Y/N**	**Y/N**	**Y/N**	**Y/N**	**Y/N**	**Y/N**	**Y/N**	**Y/N**	**Y/N**	**Y/N**
Criminal history	-.14*						-.15*			.18**	-.17**	
Attitudes							.17**				.13*	
Educational needs		.16*		.17**		.16*						
Educational assets (strength)				-.14***		-.14*						
Employment/financial		.15*	.23***	.41***	.17**	.40***						
Housing safety			.46**		.24**							
Antisocial friends			.16*	.20**		.15*						
Anger											.13*	.16*
Mental health history												
Depression (symptoms)			.25***	.23***		.20**		.15*				
Psychosis (symptoms)		.13*	.22**									
Child abuse												
Adult abuse			.14*									
Sex abuse (adult or child)												
Physical abuse (adult or child)		-.19**										
Substance abuse history			.18**	.16*		.15*						
Substance abuse (current)			.43**	.35***	.23**	.30***				.27***		.13*
PTSD[a]		.15*	.26***	.17*	.14*							
Parental difficulties			.19**				.16**			-.14*		
Parental involvement[b]		-.30**	-.29***		-.24**	-.29**				-.17*		
Relation. satisfaction (strength)				-.21**		-18**						
Family conflict							.16**					
Family support (strength)			-.21**	-.14*	-.15*							
Relationship dysfunction			.25***	.20**		.22**						
Parental stress (all)			.24**		.19**	.13*				-.14*		
Child abuse (survey)	-.15*							.14*				
Adult victimization (survey)		-.14*						-.16**				
Self-efficacy (strength)	.17**		-.22**	-.24***		-.16*						

[a]To be included in the 6-month follow-up frame, participants must be at least 3 month post interview.
[b] Base rates for Missouri arrests and convictions are extremely low, 5 and 2 percent, respectively.
***p≤.01; **p≤.05 ;*p≤.10

Table 8. Bivariate Relationship between Original WRNA Assessment Scales and 12-Month Recidivism, Missouri and Ohio.[a]

	Missouri: WRNA (N=85)						Ohio: WRNA (N=102)					
Scale	Arrests[b] Y/N	Conv.[b] Y/N	Incarc. Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N	Arrests Y/N	Conv. Y/N	Incarc. Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N
Criminal history	-.18*		.14*									
Attitudes							.14*				.14*	
Educational needs		.16*		.16*		.16*		*-.14**				
Educational assets (strength)	.17*			-.21**		-.19**						
Employment/financial	.17*	.16*	.26***	.39***	.19**	.40***	.19**	.13*		.16**	.23***	.25***
Housing safety			.37***		.15*							
Antisocial friends			.22**	.19**		.16*						
Anger	-.18**										.13*	.16*
Mental health history												
Depression (symptoms)				.23**		.21**				.14*		
Psychosis (symptoms)	-.14*			.18**		.15*				.23***		
Child abuse	.22**											
Adult abuse				.20**		.16*						
Sex abuse (adult or child)				.19**		.16*						
Physical abuse (adult or child)	.23**	-.20**										
Substance abuse history			.21**									.12*
Substance abuse (current)			.39***	.34***	.24***	.31***	.14*			.21**	.17**	.25***
PTSD[a]		.16*	.20**	.14*	.16*			.16*		.15*		
Parental difficulties			.20**					-.16**		-.14*		
Parental involvement[b]			-.31***		-.32***			-.20**		-.30***		-.19*
Relation. satisfaction (strength)				-.18**		-.16*						
Family conflict										.15*		
Family support (strength)			-.16*									
Relationship dysfunction			.22**	.19**		.21**						
Parental stress (all)			.21**	.16*		.19**		*-.20***				
Child abuse (sur.)	-.15*						.16*	.15*			.16**	.15*
Adult abuse (sur.)		-.15*		.15*		.15*	.13*					
Self-efficacy (strength)	.17*		-.31***	-.23**		-.17*				-.18**		

[a]To be included in the 12-month follow-up frame, participants must be at least 8 month post interview.
[b] Base rates for Missouri arrests and convictions are extremely low, 6 and 2 percent, respectively.
***p≤.01; **p≤.05; *p≤.10

Table 9.Bivariate Relationship between LSI-R and Original WRNA-T Assessment Scales and 6-Month Recidivism, Minnesota and Iowa.

	Minnesota: LSI-R (N=53)						Iowa: LSI-R (N=329)					
	Arrests	Conv.	Incarc.	Tech. Viol.	Offense Fail	Any Fail	Arrests	Conv.	Incarc.	Tech. Viol.	Offense Fail	Any Fail
Scale	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N
LSI-R Scales												
Criminal history	.25**			.19*	.27**	.24**		.12**		.08*	.10**	.12**
Attitudes							.07*				.08*	
Education/employment	.24**			.24**	.26**	.26**	.12***	.09*		.11**	.16***	.15***
Financial												
Accommodation					.19*							
Emotional/personal							.13**			.08*	.08*	
Leisure/recreation						.20*						
Family/marital							.16***				.15***	.08*
Antisocial friends						.20*	.09**	.07*		.11**	.11**	.13***
Substance abuse history (LSI-R)				.29**				.08*		.10**	.08*	.11**
Employment/financial (WRNA)										.11**		.11**
WRNA-T Gender-Responsive Scales												
Educational assets (strength)	-.24**			-.25**	-.25**	-.20*	-.09**				-.13**	-.11**
Housing safety							.09**	.11**		.11*	.12**	.09*
Anger	.39***	.22*		.37***	.36***	.35***	.16***			.13***	.14***	.12**
Mental health history								.13***			.12**	.10**
Depression (symptoms)												
Psychosis (symptoms)	-.21*				-.23**							
Child abuse	.29**	.18*		.31***	.24**	.28**	.09**				.11**	.07*
Adult abuse							.08*	.10**			.11**	.07*
Sex abuse (adult or child)												
Physical abuse (adult or child)	.27**	.22*		.34***	.24**	.28**	.13***	.12**		.08*	.17***	.12**
Substance abuse (current)												
PTSD[a]												
Parental difficulties						.19*		-.07*				
Parental involvement[b]	-.25*			-.36**				-.22***			-.13**	-.13**
Parental stress										.10**		.09**
Relation. satisfaction (strength)								-.14***			-.10**	
Relationship dysfunction												
Family conflict							.11**	.10**		.12***	.14***	.16**
Family support (strength)												
Child abuse (sur.)							.08*	.09**			.10**	
Adult victimization (sur.)		.19*						.10**		.09*	.10**	.12**
Self-efficacy (strength)										-.09*		-.11**

[a]To be included in the 6-month follow-up frame, participants must be at least 3 month post interview.
***p≤.01; **p≤.05;*p≤.10

Table 10.Bivariate Relationship between LSI-R and WRNA Assessment Scales and 12-Month Recidivism, Minnesota and Iowa.

	Minnesota: LSI-R (N=51)						Iowa: LSI-R (N=315)					
	Arrests	Conv.	Incar.	Tech. Viol.	Offense Fail	Any Fail	Arrests	Conv.	Incar.	Tech. Viol.	Offense Fail	Any Fail
Scale	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N
LSI-R Scales												
Criminal history	.30**	.20*		.29**	.32***	.29**	.09**	.09**		.15***	.12**	.16***
Attitudes							.14***				.11**	
Education/employment	.24**			.28**	.26**	.23*	.22***	.14***		.18***	.24***	.20***
Financial								.08*				
Accommodation		.20*		.19*	.19*	.19*				.11**		.10**
Emotional personal	.24**			.30**	.22*	.21*	.07*					
Leisure/recreation							.09*				.11**	
Family/marital				.22*			.13***				.12**	
Antisocial friends								.07*		.15***		.13***
Substance abuse history				.20*			.10**			.17***	.10**	.16***
WRNA-T, Gender-Responsive Scales												
Employment/financial										.14***		.142*
Educational strengths	-.24**				-.26**		-.13***	-.08*		-.14***	-.13***	-.14***
Housing safety												
Anger	.51***	.19*		.46***	.48***	.45***	.22***	.09**		.21***	.25***	.22***
Mental health history	.28**			.24**	.28**	.20*		.08*				
Depression (symptoms)				.19*				-.09*				
Psychosis (symptoms)		-.24**										
Child abuse	.39***	.20*		.35***	.35***	.35***						
Adult abuse												
Sex abuse (adult or child)												
Physical abuse (adult or child)	.33***			.31**	.31***	.33***	.09**	.08*		.08*	.11**	.08*
PTSD (WRNA)								-.09*			-.09*	
Parental difficulties (WRNA)	.30**			.25**	.31***	.33***						
Parental involvement (WRNA)	-.39***			-.38***	-.34**	-.28**	-.15**	-.19***			-.19***	-.13**
Parental stress(WRNA)	.23**			.23**	.19*	.21*						
Relation. satisfaction (strength)							-.07	-.10**		-.12***	-.10**	-.14***
Relationship dysfunction							.07*			.12**		.13***
Family conflict							.09*	.12**		.09**	.10**	.11**
Family support (strength)	-.18*							-.14***			-.09*	
Child abuse (sur.)				.20*								
Adult abuse (sur.)		.21*						.14***		.09**	.08*	.10**
Self-efficacy (strength)							-.10**	-.07*		-.11**	-.14***	-.13***

***p≤.01; **p≤.05; *p≤.10

Total Risk Scales-2008 Construction Validation Study

Construction validation research, completed in 2008, developed risk scores for the stand alone WRNA by summing the following risk/need scales: criminal history, antisocial attitudes, employment/financial, housing safety, anger, antisocial friends, psychotic symptoms, depression/anxiety, substance abuse history, current substance abuse, family conflict, and parental stress (see Appendix B). Strengths pertaining to educational assets, family support, and self-efficacy were subtracted from the total. It was possible to retest these scales in Ohio and Missouri. Results are shown in Table 11, below.

The findings are favorable for the Missouri site but not for the Ohio site. For the Missouri sample, the WRNA scale was predictive of returns to prison and any failure at both the 6 and 12 month follow-up time periods. Results for arrests and convictions are not shown due to limited variation on the outcome variables. It was not surprising that the results were not acceptable for the Ohio sample. As noted above, both the sampling process and the collection of follow-up data were compromised.

Results for the trailer sites, Iowa and Minnesota, are shown in Table 12. With the exception of the conviction measure, the WRNA-T showed high correlations with 12 month outcomes in the combined samples. Partial correlations were also significant, indicating that the WRNA_T offered significant incremental validity to the prediction offered by the LSI-R. Bivariate correlations between the WRNA-T alone and outcomes were stronger for the Minnesota sample than the Iowa sample. Failure to find significant incremental validity on 3 of 5 tests implicates both the small sample size and the high correlations achieved by the LSI-R. As noted above, the weak findings for the convictions tests are likely to be attributable to the low base rates resulting from the limited 12 month follow-up period. Trailer results may also be seen in Table 11, because it was possible to extract just the trailer scales from the stand alone WRNA. Results were favorable for Missouri but not Ohio.

Table 11. Predictive Validity of Stand-alone WRNA, Original 2008 Construction Validation Scales (Missouri and Ohio)

Scale	Arrests[a]		Conv.[a]		Incar.[b]		Tech. Viol.		Offense Fail		Any Fail	
	Y/N	AUC	Y/N	AUC	Y/N	AUC	Y/N	AUC	Y/N	AUC	Y/N	AUC
Revalidation Original Instrument – 6 months[c]												
WRNA Stand-alone-All (N=203)							.25***	.65	.12**	.58	.23***	.63
Levels-All (N=203)							.23***	.63	.15**	.60	.24***	.63
WRNA-T- All (N=203)							.16***		.14**		.18**	
WRNA Stand-alone -Missouri (N=91)					.40***	.83	.34***	.69	.19**	.60	.30***	.67
Levels-Missouri (N=91)					.35***	.81	.31***	.67	.15*	.60	.27***	.65
WRNA-T- Missouri (N=91)					.34***		.30***		.16*		.27***	
WRNA Stand-alone -Ohio (N=112)	--		--				--	.64	--		--	
Levels-Ohio (N=112)	.13*						--	.64	.12*	.60	.18**	.64
WRNA-T-Ohio (N-112)	--		.13*				--					
Revalidation Original Instrument - 12 Months[c]												
WRNA Stand-alone- All(N=187)							.24***	.63	.17***	.60	.24***	.63
Levels (N=187)							.24**	.63	.19***	.61	.25***	.63
WRNA-T-All (N=187)							.21***		.14***		.20***	
WRNA Stand-alone -Missouri (N=85)					.37***	.75	.30***	.68	.15*	.57	.27***	.67
Levels-Missouri (N=85)					.32***	.72	.31***	.69	.15*	.59	.28***	.68
WRNA-T-Missouri (N=85)					.27***		.27***		--		.27***	
WRNA Stand-alone -Ohio (N=102)	--		--				.16**	.64	.17**	.60	.18**	.60
Levels-Ohio (N=102)	.17**	.62	--				.16*	.65	.20**	.62	.20**	.62
WRNA-T-Ohio (N=102)	--		--				.21***		.16**		.16 *	

***p≤.01; **p≤.05; *p≤.10

[a] Arrest and conviction data are not tested for Missouri due to limited variation.

[b] Incarcerations are only tested for Missouri.

[c] Participants required at least 3 month of follow-up to be included in the 6 month data and at least 8 months to be included in the 12 month data.

Table 12. Predictive Validity of LSI-R and the Original WRNA-T, 2008 Construction Validation Scales (Minnesota and Iowa)

Scale	Arrests		Conv.		Incar.[a]		Tech. Viol.		Offense Fail		Any Fail	
	Y/N	AUC	Y/N	AUC	Y/N	AUC	Y/N	AUC	Y/N	AUC	Y/N	AUC
Revalidation Original Instrument – 6 months[b]												
Minnesota and Iowa (N=382)												
LSI-R (N=382)	.14***	.65	.12***	.66			.16***	.63	.17***	.66	.18***	.63
WRNA-T	.25***		.09*				.19***		.24***		.21***	
LSI-R + WRNA-T	.22***	.72	.13***	.68			.20***	.66	.23***	.71	.22***	.67
Partial corr.	.19***		.07*				--		.18***		.16***	
Minnesota (N=53)												
LSI-R	.23**	.65	--				.28**	.69	.22*	.63	.22**	.63
WRNA-T	.25***		--				.21**		.21**		--	
LSI-R + WRNA-T	.26**	.67	--				.27***	.69	.23**	.67	.21**	.63
Partial corr.	--		--				--		--		--	
Iowa (N=329)												
LSI-R	.14***	.69	.13***	.71			.14***	.62	.19***	.71	.18***	.64
WRNA-T	.14***		--				.14***		.15***		.16***	
LSI-R + WRNA-T	.16***	.72	.11**	.68			.16***	.63	.20***	.72	.20***	.66
Partial corr.	.08*		--				.08*		.08*		.10**	
Revalidation Original Instrument - 12 Months[b]												
Minnesota and Iowa (N=367)												
LSI-R	.21***	.68	.13***	.62			.23***	.66	.22***	.66	.22***	.64
WRNA-T	.26***		.10**				.24***		.25***		.24***	
LSI-R + WRNA-T	.27***	.70	.13***	.62			.27***	.66	.27***	.68	.26***	.67
Partial corr.	.18***		--				.15***		.17***		.16***	
Minnesota (N=51)												
LSI-R	.28***	.67	--				.33***	.68	.27**	.66	.26**	.64
WRNA-T	.36***		--				.31***		.33***		.27***	
LSI-R + WRNA-T	.34***	.68	--				.35***	.69	.32***	.66	.29**	.64
Partial corr.	.24**		--				--		.20**		--	
Iowa (N=316)												
LSI-R	.22***	.69	.14***	.65			.22***	.66	.22***	.68	.22***	.65
WRNA-T	.16***		--				.19***		.18***		.20***	
LSI-R + WRNA-T	.23***	.70	.13***	.63			.25***	.66	.24***	.68	.25***	.68
Partial corr.	.08*		--				.11**		.09*		.12**	

***p≤.01; **p≤.05 ;*p≤.10

[a] Incarcerations are only tested for Missouri.

[b] Participants required at least 3 month of follow-up to be included in the 6 month data and at least 8 months to be included in the 12 month data.

Revision of the WRNA Scales

With the exception of the Ohio results, the results shown above were within the findings expected for a 12 month recidivism study and many individual scales showed adequate relationships to outcome measures. There were nevertheless some concerns for the following scales: criminal history, family conflict, housing safety, and the relationship scales. We also sought to create a shorter instrument by testing interview alternatives for several of the survey scales, e.g., relationships, parental stress, adult abuse, and child abuse. Finally, a number of test items were recommended by practitioners and UC researchers to improve scales that already had sufficient predictive validity.

Improvement of the assessment scales was achieved through the development and testing of new items on a new construction validation sample and then revalidating those items. It was hoped that the revalidation of the new scales would reduce concerns for the need of another revalidation study. However, the number of cases available for testing depended upon whether or not the scale was a gender-neutral or a gender-responsive scale. Gender-neutral scales were only tested in the 2 sites that examined the stand- alone assessment. Therefore, tests of criminal history, antisocial attitudes, educational needs, antisocial friends, and substance abuse scales were only tested in Ohio and Missouri. Sample sizes for the construction sample and revalidation samples consisted of 101 and 102 participants, respectively at 6 months, and 94 and 93 at 12 months. These sample sizes were quite small for these purposes.

Tests of the gender-responsive scales utilized more favorable numbers. At 6 months the construction sample and revalidation sample included 292 and 293 participants, respectively. At 12 months, tests of the gender-responsive scales involved 276 participants in the construction sample and 277 in the revalidation sample.

Because these numbers are somewhat smaller than desired, especially for the gender-neutral variables a second examination of revised scales was made for each of the research sites. Given the sample-specific nature of some findings, we retained scales that may not have reached adequate predictive validity in the revalidation sample, but did in the test sites. Results for the construction and revalidation samples are shown in Tables 13 and 14. State

specific findings appear in Tables 15 and 16 for the stand-alone WRNA sites (Missouri and Ohio) and in Tables 17 and 18 for the WRNA-T trailer sites (Iowa and Minnesota). LSI-R domain scales are not the subject of these analyses. Alpha measures of internal consistency are shown in Appendix E. A discussion of changes made to specific scales follows.

Gender Neutral Scales:

CRIMINAL HISTORY: Problems with the original criminal history scale were known in advance, because findings were also observed to be rather weak during the 2008 construction validation study. 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 ones that contributed to the predictive validity of the scale were:

1. Was your last conviction (felony or misdemeanor) within the past three years?
2. Age at intake
   18-30=2
   31-40=1
   41+=0

As shown in Table 13 and 14, these changes improved the scale considerably for construction and revalidation samples. State-specific finding show that the improvements were primarily attributable to the Missouri sample, however.

The revised assessment scale also mandates use of official records of prior offenses. In most sites, this study had access to both self-report and official accounts. Discrepancies were detected, and the official records proved to be more valid and were substituted for self-report wherever possible. Alpha for the scale was low (.46), which is not unusual for a criminal history scale containing rather diverse items. The revised criminal history scale was included in the final stand-alone WRNA.

ATTITUDES: The attitudes or criminal thinking scale was related to offense-related outcomes in an inconsistent manner. Yet, it demonstrated a high alpha =.83 reliability rating. Item analysis revealed that most of the items were not predictive of any of the outcomes for the combined Ohio and Missouri samples. The fact that the scale could not be improved upon may implicate the underlying construct as not as meaningful as other cognitive issues among women offenders. With respect to cognitive patterns, anger and self-efficacy offered far better contributions to the predictive validity of the

Table 13. Bivariate Relationship between Revised WRNA Scales and 6-Month Recidivism, Construction and Revalidation Samples.

	Construction Sample						Revalidation Sample					
Scale	Arrests[a] Y/N	Conv.[a] Y/N	Incar.[b] Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N	Arrests[a] Y/N	Conv.[a] Y/N	Incar.[b] Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N
WRNA Gender-Neutral (N=101, N=102)												
Criminal history			.49	.28***	.23***	.20***			.49***	.28***	--	.19**
Antisocial friends			.22*								-.14*	
Substance abuse history				.29***		.30***			.32***	.26***		.16*
Current substance abuse			.24*	.25***		.22**			.64***	.45***	.25***	.35***
Gender-Responsive Risk Factors (N=292, N=293)												
Employment/financial (N=101,102)				.25***	.20**	.33***			.26**	.35***		.31***
Depression collapsed			.29**					.16***	.29**		.08*	.09*
Family conflict (101,102)					.21**						.21***	
Family support (collapsed)	-.13**	-.09*	-.25**		-.16***				-.23*			
Parental involvement (N=181;N=176)	-.14**	-.23***	-.44***	-.19***	-.23***	-.19***			-.25*	-.15**	-.15**	-.15**
Parental stress (all)-Collapsed				.09*					.34***	.09*	.15***	.13*
Relationship dysfunction-old									.33***			
Relationship dysfunction-new									.23*			.13**
Relationship satisfaction				-.11*						-.08*		
Self-efficacy (collapsed)	-.27***	-.10***	-.20*	-.24***	-.21***	-.27***	-.11**	-.14**	-.21*	-.14	-.10**	-.13**

[a]Missouri cases are omitted from tests of arrest and conviction outcomes. Shaded area indicates the test was not conducted.
[b]Incarceration data are only available for the Missouri cases.
***p≤.01; **p≤.05; *p≤.10

Table 14. Bivariate Relationship between Revised WRNA Scales and 12-Month Recidivism, Construction and Revalidation Samples.

	Construction Sample						Revalidation Sample					
Scale	Arrests[a] Y/N	Conv.[a] Y/N	Incar.[b] Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N	Arrests[a] Y/N	Conv.[a] Y/N	Incar.[b] Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N
WRNA Gender-Neutral (N=94, N=93)												
Criminal history			.36***	.26***		.19**			.55***	.22**	.17**	.17**
Antisocial friends			.34**	.15*	.24***	.24***					-.14*	
Substance abuse history				.25***	.20**	.34***			.32***	.24***		
Current substance abuse			.28**	.28***	.22**	.37***			.57***	.37***	.27***	.28***
Gender-Responsive Risk Factors (N=276, N=277)												
Employment/financial (N=94,93)			.20*	.31***	.30***	.41***			.27**	.35***	.14*	.32***
Depression collapsed										.08*		
Family conflict (94,93)											.16**	
Family support – collapsed	-.13**	-.10*	-.30**		-.13**	-.08*	-.12**	-.10*			-.12**	
Parental involvement (N=170, N=165)	-.20***	-.24***	-.45***		-.27***	-.26***	-.21***	-.18**	-.34**	-.11*	-.22***	-.14**
Parental stress (all) - collapsed	.09*					.08*			.44***	.11**		.10*
Relationship dysfunction – old	.12**		.21*			.11*			.26**			
Relationship dysfunction - new	.12**				.08*	.13**						
Relationship support	-.12**				-.16***	-.20***				-.08*		-.08*
Self-efficacy – collapsed	-.20**		-.35***		-.21***	-.19***	-.13**	-.12**		-.17**	-.14***	-.15***

[a]Missouri cases are omitted from tests of arrest and conviction outcomes. Shaded area indicates the test was not conducted.
[b]Incarceration data are only available for the Missouri cases.
***p≤.01; **p≤.05 ;*p≤.10

Table 15. Bivariate Relationship between Revised WRNA Scales and 6-Month Recidivism, Missouri and Ohio Samples.

	Missouri (N=91)						Ohio (N=112)					
Scale	Arrests[a] Y/N	Conv. Y/N	Incar.[b] Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N	Arrests[a] Y/N	Conv.[a] Y/N	Incar.[b] Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N
WRNA Gender-Neutral												
Criminal history			.47***	.21**	.32***	.15***				.18**		
Antisocial friends			.15*	.18**		.14*						
Substance abuse history			.19**	.17**		.16**						
Current substance abuse			.45***	.35***	.27***	.31**				.22***		
Gender-Responsive Risk Factors (N=292, N=293)												
Employment/financial			.22**	.40***	.17**	.40***						.13*
Depression – collapsed			.28**	.16*	.15*	.15*						
Family conflict							.19**	.15*			.21***	.17**
Family support – collapsed			-.24***		-.18**			-.13*				
Parental involvement (N=181, N=176)			-.35***	-.21**	-.31***	-.23**				-.17*		
Parental stress (all) - collapsed			.21**		.19**	.23**						
Relationship dysfunction – old			.25***	.19**		.22**						
Relationship dysfunction - new			.20**	.17**		.19**						
Relationship satisfaction				-.19**								
Self-efficacy – collapsed			-.20*	-.23***		-.16*				.15*		

[a]Missouri cases are omitted from tests of arrest and conviction outcomes. Shaded area indicates the test was not conducted.
[b]Incarceration data are only available for the Missouri cases.
***p≤.01; **p≤.05 ;*p≤.10

Table 16. Bivariate Relationship between Revised WRNA Scales and 12-Month Recidivism, Missouri and Ohio Samples.

	Missouri (N=85)						Ohio (N=102)					
	Arrests[a]	Conv.[a]	Incar.[b]	Tech. Viol.	Offense Fail	Any Fail	Arrests[a]	Conv.[a]	Incar.[b]	Tech. Viol.	Offense Fail	Any Fail
Scale	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N
WRNA Gender-Neutral (N=101, N=102)												
Criminal history			.44***	.17*	.22***							
Antisocial friends			.22**	.19**		.16*						
Substance abuse history			.23**									
Current substance abuse			.41***	.35***	.27***	.32**	.16*			.18**	.18**	.25***
Gender-Responsive Risk Factors (N=292, N=293)												
Employment/financial (N=101,102)			.23**	.38***	.17*	.40***	.21**	.18**		.19**	.24***	.28***
Depression – collapsed			.15*	.16*		.16*						
Family conflict (N=101,102)					.18**					.16**		
Family support – collapsed			-.26**							-.13*		
Parental involvement (N=181, 176)			-.40***		-.25**	-.20*		-.18*		-.31***		-.21**
Parental stress (all) - collapsed			.20**	.15*		.22**		-.16**				
Relationship dysfunction-old			.22**	.18**		.21**						
Relationship dysfunction-new			.14*	.18**		.20**						
Relationship satisfaction				-.15								
Self-efficacy – collapsed			-.27***	-.17*								

[a]Missouri cases are omitted from tests of arrest and conviction outcomes. Shaded area indicates the test was not conducted.
[b]Incarceration data are only available for the Missouri cases.
***p≤.01; **p≤.05; *p≤.10

Table 17. Bivariate Relationship between Revised WRNA Scales and 6-Month Recidivism, Iowa and Minnesota Samples.

	Iowa (N=329)						Minnesota (N=53)					
	Arrests[a]	Conv.[a]	Incar.[b]	Tech. Viol.	Offense Fail	Any Fail	Arrests[a]	Conv.[a]	Incar.[b]	Tech. Viol.	Offense Fail	Any Fail
Scale	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N
Employment/financial				.10**		.09**						.13*
Depression collapsed												
Abuse - all items	.11**	.09*			.14***	.09**						
Family conflict	.13***				.14***	.16***	.19**	.15*			.18***	.17**
Family support – collapsed								-.13*				
Parental involvement		-.24***			-.16**	-.14**				-.32**		
Parental stress (all) - collapsed				.10**		.10**						
Relationship dysfunction – old												
Relationship dysfunction - new						.08*						
Relationship satisfaction		-.11**			-.08*							
Self-efficacy – collapsed	-.09*	-.13***		-.12***	-.11**	-.16***				-.21*		

[a]Missouri cases are omitted from tests of arrest and conviction outcomes. Shaded area indicates the test was not conducted.
[b]Incarceration data are only available for the Missouri cases.
***p≤.01; **p≤.05; *p≤.10

Table 18. Bivariate Relationship between Revised WRNA Scales and 12-Month Recidivism, Iowa and Minnesota Samples.

	Iowa (N=316)						Minnesota (N=51)					
Scale	Arrests[a] Y/N	Conv[a] Y/N	Incar.[b] Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N	Arrests[a] Y/N	Conv.[a] Y/N	Incar.[b] Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N
Employment/financial				.13***		.12**						
Depression (collapsed)							.21*			.29**	.22*	.24*
Family conflict (101,102)	.11**	.09**			.11**	.14***						
Family support (collapsed)	-.08*	-.12**			-.10**		-.29**			-.25**	-.25**	-.22*
Parental involvement	-.14**	-.20***			-.20***	-.14**	-.36**			-.32**	-.32**	-.24*
Parental stress (all)-Collapsed						.08*		.27**		.22*	.22*	.20*
Relationship dysfunction-old	.07*			.12**		.13***						
Relationship dysfunction-new	.08*			.11**		.13***						
Relationship satisfaction	-.08*	-.09***		-.12**	-.10**	-.13***						
Self-efficacy (collapsed)	-.11**	-.10**		-.14***	-.17***	-.16***						

[a]Missouri cases are omitted from tests of arrest and conviction outcomes. Shaded area indicates the test was not conducted.
[b]Incarceration data are only available for the Missouri cases.
***p≤.01; **p≤.05 ; *p≤.10

total scale. The attitudes scale is included in the needs section of the assessment (Part IV) but is not included on the risk scale.

EDUCATIONAL SCALES: The educational needs scale and the educational strengths scale wererelated to offender outcomes for the sample as a whole. However, correlations were primarily related to technical violations and to the summary item capturing any failures. The items comprising each scale were standard educational items, and no test items were introduced to improve the scales.

> Item analysis revealed that the item capturing completion of high school or the GED detracted from the predictive validity of the **Educational Needs** scale. Items pertaining to reading difficulties, special needs, and special education were far more predictive. Exclusion of the item failed to improve the predictive validity of the scale, or its reliability. As a result, GED/high school completion was retained on the scale (alpha=.70).
>
> Because, item analysis of the **Educational Strengths** scale revealed that all of the items contributed to the scale, no changes were made (alpha=.63).

ANTISOCIAL FRIENDS: The Antisocial Friends scale was only predictive of returns to prison for the Missouri sample. Upon further analysis, one item was found to be detracting from the predictive validity of the scale: "Prior to your arrest, did you have some friends who seemed supportive of you?" Its omission improved the predictive validity for the construction but not the revalidation sample. It also made no change to the predictive validity in the state samples. Alpha was equal to .72.

SUBSTANCE ABUSE: Two substance abuse scales were created for the Women's Risk/Needs Assessments--substance abuse history and current/recent substance abuse. Although they both were valid in most samples, the present study examined the addition of two questions, one for each scale.

> **SUBSTANCE ABUSE HISTORY**: The addition of an item capturing previous use of opiates, hallucinogens or ecstasy improved the predictive validity of the scale especially during the 12 month follow-up period. Alpha for the scale was substantial=.89.
>
> **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 scale in the construction and validation sites and for both the Missouri and the Ohio samples. Alpha for the new scale was .76.

Gender-Responsive Scales:

EMPLOYMENT/FINANCIAL: The employment/financial scale was a robust predictor of outcomes in most community samples. Even so, a number of items were tested to determine whether the scale could be improved. Interviewers voiced concern that the original items were more appropriate to middle class samples than to samples of justice-involved women, e.g., do you have a checkbook, do you own an automobile, do you have a savings account. These items, however, continued to predict in the new probation samples.[12] We also tested a number of items that were suggested by practitioners and administrators in the study sites. Among the test items, two were found to contribute to the employment financial scale: Those making substantial contributions were:

1. Do you make less than $10,000 per year?
2. Do you live in a household where at least one member has full-time, year-round employment?

The test items, however, were not collected at the trailer sites. This necessitated creating the new scale on the data from the stand alone sites and using the original scale when testing the cumulative WRNA-T. When the cumulative scale, assembling all of the risk/needs scales, was created, it was necessary to prorate the old 8 point scale to a 10 point scale in order to be comparable to the Ohio and Missouri scale. Alpha for this scale was low (.55). However, the low alpha was trumped by fairly high predictive validity of the scale.

HOUSING SAFETY: The original housing safety scale was significantly related to a number of outcomes. This study examined 5 new interview items in an attempt to determine whether they could improve the predictive validity of the scale. Item analysis revealed that the scale could not be improved. Alpha for the original scale was .70

ANGER: The original scale offered predictions to outcomes in all accept the Missouri sample, and item analysis failed to indicate any interview questions that were not contributing to the scale. Therefore, no changes were made (alpha=.71).

HISTORY OF MENTAL ILLNESS: The History of Mental Illness scale was only predictive of offense-related outcomes in Minnesota and to a modest degree in Iowa.

[12] This was only true for the probation assessment. The pre-release and institutional scales did not benefit from items pertaining to checkbooks, saving account and automobiles.

An attempt to improve the scale by omitting a question about whether the participant had ever attempted suicide failed to improve the scale. The scale was retained in its original form (Alpha=.81) but was not included in the cumulative risk scale.

CURRENT SYMPTOMS OF MENTAL ILLNESS: Predictions for current symptoms of mental illness tended to be sample specific. For the total samples, the depression scale predicted to some 6 month outcomes, and both items (depression and psychosis) were highly predictive of incarcerations at 6 months in the Missouri sample.

> **DEPRESSION:** The original WRNA Depression Scale was weakly correlated with some outcomes. Item analysis revealed this to be true of most of the items as well. As a result no changes were made to the actual scale. Alpha for the cumulative scale was .73. Collapsing the scale into a 3 point scale, consistent with other studies, improved its predictive validity in the construction and revalidation samples at 6 months.
>
> **PSYCHOTIC SYMPTOMS**: This scale was comprised of two items, which generally did not predict well except in the Missouri sample. There were no test variables to make any amendments to the scale. The inter-item correlation (alpha was inappropriate) was $r=.15, p \leq .001$.

ABUSE-INTERVIEW SCALES: The interview furnished four questions that enabled the creation of 4 abuse scales: a) adult victimization; b) child abuse; c) sex abuse (experienced by an adult or a child); d) physical abuse (experienced as an adult or a child). When samples were combined, these items were predictive of some of the outcome variables, especially the child abuse and physical abuse scales. With only the original questions asked, there was no possibility of modifying the scales. 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. Inter-item correlations were as follows: a) child abuse ($r=.44, p \leq .001$); b) adult abuse ($r=.35, p \leq .001$); sexual abuse ($r=.31, p \leq .001$); physical abuse ($r=.23, p \leq .001$).Two of the four variables were used in the final risk/needs assessment (child abuse and adult abuse).

PTSD: Four interview items were based upon the Veteran's Administration's Post Traumatic Stress Disorder Scale. The cumulated scale was predictive in the Missouri sample. Alpha for the scale was .82. It was not included in the risk scale but appears in section IV of the assessment.

FAMILY OF ORIGIN SCALES: Two scales were created, one measuring Family Conflict and another Family Support. As can be seen in Tables 5 and 6, correlations with the outcome measures were modest. Results are adversely affected by

confusion among interviewers at the Ohio site who reported that many participants had difficulties conceptualizing siblings and parents in contexts of blended families, separated families, parent figures, etc. Thus, the revised interview is restructured to capture a more nuanced definition of family of origin. Measures to resolve the confusion will also be incorporated into the interviewers' training curriculum.

FAMILY CONFLICT: One test item provided a means of making modest improvements to the scales. The question inquired about whether parents or siblings tended to be critical of the participant. The item was not available for the Iowa or the Minnesota samples. For uniformity in computing the final family conflict scale, the 3 point Iowa and Minnesota family conflict scales were prorated to a 4 four point scale. The scale alpha was low .43.

FAMILY SUPPORT: No test variables were entered into the study to allow for modification of this scale. As with other studies, however, collapsing the scale improved its validity somewhat (low=0)(medium=1-3)(high=4). Alpha for this scale was .73

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. The earlier construction validation study used only the survey parental stress scale, and did so with favorable results. The goal in the present study was to determine whether an interview scale of the same would suffice.

PARENTAL INVOLVEMENT: This interview scale was not included in the cumulative risk/needs scale because it was valid only when mothers responded. In other words, non-parents could not be included as 0 on the scale. A test item indicating whether or not the participant was having difficulty maintaining or obtaining custody was determined to be relevant to parental involvement scale. In this case an answer of "no" was indicative of involvement. When this item was added to the scale, improvements in predictive validity were seen in both the construction and validation samples. The item was a potent predictor in all of the samples. Alpha was .66 for the revised scale.

PARENTAL DIFFICULTIES: The attempt to create an interview scale to substitute for the survey scale was not successful. Even in cases where this scale was predictive, the survey scale (below) was superior.

PARENTAL STRESS: Correlations were observed for the entire sample, and no items were observed to be detracting from the scale. There are however, 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 (13). In this sample, 56 women (15.5 percent of the women with children under 18) did not answer at least one of the scale questions.
3. Once the first two scoring steps have been completed, non-parents are entered into the scale as 0.

When this scale is added to the total risk/needs scale, it was collapsed into high, medium and low values:

Low (0) = 0-9
Medium (1) = 10-18
High (2) = 19+

Results were significant in Missouri and Iowa. Alpha was high (.83).

INTIMATE RELATIONSHIP SCALES: For a number of reasons, the original relationship scales, relationship satisfaction, and relationship dysfunction, seldom correlated with post-release outcomes. 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, even when they were not case management or test variables. 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.

Even so, item analysis revealed ways that the scales could be reconstructed. To reduce the possibility of interviewer bias and improve the privacy of the questions, data for relationship dysfunction questions will be collected in the survey portion of the assessment. The interview will collect information on relationship support. The scales are as follows:

RELATIONSHIP DYSFUNCTION: The items comprising this scale are as follows:

1. In general would you describe these relationships as supportive and satisfying?
2. Do you get into relationships that are painful for you? Or is your present relationship a painful one?
3. Have significant others loved and appreciated you for who you are?
4. 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?
5. Do you tend to get so focused on your partner that you neglect other relationships and responsibilities?
6. Have partners been able to convince you to get involved in criminal behavior?

The fourth item detracted from the predictive validity of the scale. Even with its removal, however, the scale made few predictions to outcome and alpha was low (.50).

RELATIONSHIP SATISFACTION: Probation participants involved with significant others tended to evidence lower rates of recidivism. The first three items on the original interview spoke to involvement and satisfaction in an intimate relationship. These were added to a fourth item from the interview (Alpha was equal to .81. They are as follows).

1. Are you involved with a significant other?
2. Are you married?
3. Have you been involved with this person for 6 years or more?

4. Is your current relationship satisfying to you (i.e., does it make you happy at the present time? [If no significant other, indicate no.]

Neither revised scale is included in the summed risk scale. Both appear in Part IV of the revised assessment.

SELF-EFFICACY: The self-efficacy was a well-established Rosenberg Self-Efficacy scale (Alpha=.87) that we did not wish to make improvements to. For addition to the cumulative risk/needs scale, the scale was collapsed into high (24+) and low (0-23) values.

CHILD ABUSE and ADULT ABUSE SURVEY SCALES: The validity of these survey scales appeared to vary by sample. The findings were modest and failed to improve upon results found in for the interview scale. As with the pre-release tool, we made the decision to omit these items from further assessments. The items were uncomfortable for some interviewers and participants and in the latest round of research, they provide no more than what could be obtained through the interview. Their omission also provided a valuable opportunity to shorten the overall assessment

Revision of the Total Risk Scales

Compilation of revised WRNA cumulative scales followed a process similar to that used for the individual scales. Scales were first developed for a construction validation sample and then retested on a revalidation sample. We begin with the construction of the WRNA-T (see Tables 19 and 20) and then move to the construction of a stand-alone assessment (see Table 21 and 22).

Construction of the Revised WRNA-T

Through analysis of the construction sample, the optimal items for inclusion in the WRNA-T cumulative scale consisted of the following separate risk/needs scales: Employment/financial, housing safety, anger, depression (collapsed), child abuse, adult abuse and parental stress. The scale also involved subtracting scores for strengths pertaining to educational assets, self-efficacy and family support. The scale formed for the construction sample proved to be stable in the new tests on the revalidation sample Results for the construction and validation samples are shown in Table 19, below. Scale refinements discussed in the previous section represent an improvement to the predictive validity of the

trailer in comparison to results shown in Tables 11 and 12 above. The 2008 trailer was showed sufficient revalidation results even prior to the attempted refinements. Even so, the cumulative scale, showed stability over the construction and revalidation samples. Moreover, the revised trailer cumulative scale contain items pertaining to abuse (child and adult), which did not appear in the original trailer. The revised scale also does not include a risk factor identifying psychotic behaviors. This was seldom predictive and is included in Part IV of the new assessment.

An additional examination apart from the construction and validation samples is shown in Table 20. Here results are examined for the LSI-R/trailer sites (Iowa and Minnesota). Table 20 shows some improvement in the predictive validity of the revised WRNA-T over the original (results which are shown in Table 12). However, the improvement is largely attributable to the Minnesota results. Results for Iowa are remarkably similar to those shown for the 2008 trailer. In addition, the overall findings for Iowa are less favorable than those for Minnesota.

In a state-specific report prepared for Iowa officials, we speculated that the attenuated findings could be attributable to the fact that a large portion of the Iowa participants were engaged in a demonstration project testing the Women Offender Case Management Model (WOCMM) (Van Dieten, 2008). Over the year following their interviews, they were engaged in intensive programming targeted to gender-responsive risk factors. If successful, the programming could have attenuated the impact of these variables. In fact when the analysis disaggregated the sample and examined results for the non-WOCMM participants, the results were much more favorable for the non-WOCMM participants (see Van Voorhis, Bauman, & Brushett, 2012).

Overall, however, the findings speak favorably to the predictive validity of the trailer. It offered statistically significant contributions to the LSI-R in 26 of 30 tests. The exceptions involved the convictions outcome measure which, due to the limited follow-up time frame evidenced low base rates known to compromise statistical tests. For the total

Table 19. Bivariate Relationship between Revised WRNA Trailer (WRNA-T) and Outcomes for Construction and Revalidation Samples, All Sites.

	Construction (N=276)						Revalidation (N=277)					
Scale	Arrests[a] Y/N	Conv.[a] Y/N	Incar.[b] Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N	Arrests[a] Y/N	Conv[a]. Y/N	Incar. [b] Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N
6 Month Follow-up (N=292, N=293)												
WRNA_T	.25***	.10*	--	.18***	.22***	.23***	.27***	.13**	.52***	.26***	.27***	.26***
12 Month Follow-up (N=276, N=277)												
WRNA-T	.25***	.11**	--	.24***	.24***	.25***	.26***	.17***	.46***	.28***	.26***	.26***

[a] Analysis of arrest and conviction data excluded Missouri cases, because of limited variation. . Analysis at 6 months included 248 in the construction sample and 246 in the revalidation sample. Analysis at 12 month included 236 in the construction sample and 232 in the revalidation sample.

[b] Analysis is only for Missouri cases. Analysis at 6 months included 44 in the construction sample and 47 in the revalidation sample. Analysis at 12 month included 40 in the construction sample and 45 in the revalidation sample.

***p≤.01; **p≤.05 ;*p≤.10

Table 20. Bivariate Relationship between LSI-R and Revised WRNA Trailer (WRNA-T and Outcomes for Iowa and Minnesota.

	Arrests		Conv.		Incar.[b]	Tech. Viol.		Offense Fail		Any Fail	
Scale	**Y/N**	**AUC**	**Y/N**	**AUC**	**Y/N**	**Y/N**	**AUC**	**Y/N**	**AUC**	**Y/N**	**AUC**
6 Month Follow-up											
Total (N=382)											
LSI-R	.14***	.65	.12***	.66		.16***	.63	.17***	.66	.18***	.63
WRNA-T	.30***		.12***			.22***		.29***		.25***	
LSI-R+WRNA-T	.24***	.73	.14**	.69		.22***	.67	.26***	.73	.24***	.68
Partial corr.	.26***		.07*			.17***		.24***		.19***	
Iowa (N=329)											
LSI-R	.14***	.69	.13***	.71		.14***	.62	.19***	.71	.18***	.64
WRNA-T	.16***		--			.14***		.18***		.18***	
LSI-R+WRNA-T	.17***	.72	.12**	.70		.16***	.64	.21***	.73	.21***	.67
Partial corr.	.10**		--			.09*		.10**		.11**	
Minnesota (N=53)											
LSI-R	.23**	.65	--	.60		.28***	.69	.22*	.63	.22*	.63
WRNA-T	.36***		--			.34***		.32***		.28***	
LSI-R+WRNA-T	.30***	.70	--	.61		.33***	.73	.28**	.68	.26**	.67
Partial corr.	.28***		--			.21*		.23**		.18*	
12 Month Follow-up											
Total (N=366)											
LSI-R	.21***	.68	.13***	.62		.23***	.66	.22***	.66	.22***	.64
WRNA-T	.29***		.16***			.26***		.29***		.26***	
LSI-R+WRNA-T	.28***	.71	.16***	.65		.28***	.67	.29***	.70	.28***	.67
Partial corr.	.22***		.11***			.18***		.22***		.19***	
Iowa (N=315)											
LSI-R	.22***	.70	.14***	.65		.22***	.66	.22***	.68	.22***	.65
WRNA-T	.18***		.10**			.21***		.20***		.21***	
LSI-R+WRNA-T	.23***	.70	.14***	.65		.25***	.66	.25***	.69	.25***	.66
Partial corr.	.09**		--			.12**		.11**		.12***	
Minnesota (N=51)											
LSI-R	.28**	.67	--	.58		.33***	.69	.27**	.66	.26**	.64
WRNA-T	.44***		--			.37***		.40***		.36***	
LSI-R+WRNA-T	.37***	.71	--	.59		.37***	.71	.35***	.69	.32***	.67
Partial corr.	.36***		--			.21*		.31***		.25**	

***p≤.01; **p≤.05 ;*p≤.10

sample and the individual sites, AUC values equaled or surpassed .70 on at least one of the follow-up measures tested per time period and location.

Construction of the Stand-Alone WRNA

The optimal complete WRNA cumulative scale consisted of the following individual risk/need scales:

Criminal history
Antisocial friends
Substance abuse history
Current substance abuse
Depression (collapsed)
Employment/financial
Housing safety
Anger
Child abuse
Adult abuse
Parental stress

The following strengths are subtracted:

Educational assets
Self-efficacy (collapsed)
Family support (collapsed)

Results for the construction and reconstruction samples are shown in Table 21. Predictive validity for the total scale was strong for measures of incarceration (Missouri), technical violations, and any failure, at the 12 month follow-up point. Results found in the construction sample were confirmed by the revalidation sample.

State specific findings appear in Table 22. Results generally show an improvement over the 2008 assessments. Nevertheless, these findings do vary by state and outcome measure. Predictions of returns to prison, technical violations, offense-related failures, and any type of failure are strong for the Missouri sample. Correlations are high, and AUC values equal or approach .70. However, results for the Ohio sample are unacceptable and indicative of errors in the sample selection process and/or the collection of follow-up data.

Table 21. Bivariate Relationship between Revised WRNA Stand-Alone and Outcomes for Construction and Revalidation Samples, All Sites.

	Construction						Revalidation					
Scale	Arrests[a] Y/N	Conv.[a] Y/N	Incar.[b] Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N	Arrests Y/N	Conv. Y/N	Incar.[b] Y/N	Tech. Viol. Y/N	Offense Fail Y/N	Any Fail Y/N
6 Month Follow-up (N=101, N=102)												
WRNA-Stand Alone	--	--	.34**	.31***	.14*	.32***	--	--	.60***	.40***	.18**	.29***
12 Month Follow-up (N=94, N=93)												
WRNA-Stand Alone	.26***	--	.39***	.35***	.27***	.42***	--	--	.57***	.32***	.16*	.19**

[a] Analysis of arrest and conviction data excluded Missouri cases, because of limited variation. Analysis at 6 months included 57 in the construction sample and 55 in the revalidation sample. Analysis at 12 month included 54 in the construction sample and 48 in the revalidation sample.

[b] Analysis is only for Missouri cases. Analysis at 6 months included 44 in the construction sample and 47 in the revalidation sample. Analysis at 12 month included 40 in the construction sample and 45 in the revalidation sample.

***p≤.01; **p≤.05; *p≤.10

Table 22. Bivariate Relationship between the WRNA Stand-Alone Assessment and Outcomes For Missouri and Ohio.

	Arrests		Conv.		Incar.[b]		Tech.Viol.		Offense Fail		Any Fail	
Scale	**Y/N**	**AUC**	**Y/N**	**AUC**	**Y/N**	**AUC**	**Y/N**	**AUC**	**Y/N**	**AUC**	**Y/N**	**AUC**
6 Month Follow-up												
Total (N=203)												
WRNA							.35***	.71	.16***	.59	.31***	.67
Levels							.36***	.70	.13**	.58	.29***	.65
WRNA-T							.20***		.15**		.22***	
Missouri (N=91)												
WRNA					.48***	.89	.37***	.71	.26***	.64	.33***	.68
Levels					.41***	.82	.38***	.70	.24***	.64	.33***	.68
WRNA-T					.34***		.30***		.16***		.29***	
Ohio (N=112)												
WNRA	--		--				.13*	.67	--	--	--	--
Levels	--		--				.12*	.67	--	--	--	--
WRNA-T	--		--				--		--	--	--	--
12 Month Follow-up												
Total (N=187)												
WRNA							.34***	.69	.22***	.62	.31***	.67
Levels							.35***	.69	.22***	.62	.31***	.66
WRNA-T							.24***		.17***		.24***	
Missouri (N=85)												
WRNA					.46***	.80	.32***	.70	.24**	.62	.28***	.67
Levels					.39***	.76	.36***	.69	.25***	.64	.29***	.68
WRNA-T					.28***		.28***		--		.27***	
Ohio (N=102)												
WRNA	--	--	--	--			.17**	.65	.16**	.60	.20**	.62
Levels	--	--	--	--			.16**	.63	.14*	.58	.16**	.59
WRNA-T							.17**		.18**		.18**	

***p≤.01; **p≤.05; *p≤.10

The revised assessments appear in Appendices F through I.

CONCLUSION

Validation tests of both the 2008 WRNA and WRNA-T as well as the revised instruments produced acceptable results that afford a good deal of confidence in these tools. 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. The Ohio findings stand in contrast to the generally favorable findings for the other three sites, however. As noted earlier, the study, which was voluntary, was terminated for lack of participation on the part of probation officers and potential participants.

The larger 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 avoided 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 validity of the LSI-R as a prediction was favorable (see Table 20).

We did not succeed in a goal of creating a single tool for use with pre-release and probation settings. The probation tool presented in this reportdiffers primarily with respect to the construction of the employment/financial scale and the composition of the total risk scale.

Specifically, the prerelease tool fits a more troubled population. More mental health scales are included in the risk scale for the prerelease tool, for example, than for the probation tool. Moreover, the employment/financial scale of the pre-release tool taps indicators of more entrenched poverty than the probation tool. Simply put, creating a single tool could have compromised the validity of each one.

Notwithstanding these contributions, there are some necessary precautions to be taken in understanding these findings. The more ideal research sample would have involved a random statewide sample, or several of them. Two of the three sites sampled for this study, truncated assessment distributions, through a process which attempted to screen-out low risk women. A third probation site was affected by poor cooperation from probation officers in referring women to the study. Refusal rates were not high once probationers met with UC research staff, however, they could have been very high during probation officer presentations of the study to prospective probationers. The one site (Missouri) which tapped all potential, English-speaking clients was delayed in starting, a fact which reduced size of the sample for that site.

Follow-up data are also likely to be truncated. The follow-up time period for the present study was 12 rather than 24 months. The earlier 2004 - 2008 studies found more impressive results at 24 months than at 12, and 24 months is the recommended follow-up observation period for both program evaluations and prediction studies. Limited base rates are known to attenuate findings, and longer follow-up periods improve base rates, which in turn tends to improve predictive validity coefficients. Of greatest concern in this regard involves mental health scales. In other studies, these often did not appear to emerge as correlates until the 18 to 24 month time frame.

With the exception of Iowa, where the assessment was used for case planning for many women offenders, the study samples are rather small. Sufficient statistical power for a study where construction and revalidation samples are intended would typically require about 800 cases. The present study amassed data on 203 cases for gender-neutral variables and 585 cases for gender-responsive variables. This necessitated a boot-strap approach where scales were developed in a construction validation sample and retested in a revalidation sample as well as in state-specific samples.

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 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, in some tests results for the LSI-R trailer were not as favorable as those for the WRNA stand-alone instrument. Tables 12 and 20 show a number of instances where the validity of the WRNA-T, found to evidence acceptable predictive validity on its own, became more limited when added to the LSI-R scores. That is validity was "pulled down" by the LSI-R rather than the other way around. We note that interviewers for the WRNA assessments were trained immediately prior to data collection. In contrast a number of state officials observed that many of the LSI-R interviewers were due to be retrained. 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.

Even with these limitations, results are somewhat more favorable than typically seen at a 12 month follow-up. It is likely that the study limitations did not bias findings in a favorable direction. Typically, limitations with base rates, sample size, and quality assurance attenuate findings. Most importantly, we have much more confidence in the stability of the assessment, because it now is seen to be predictive across several jurisdictions.

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 practiceand guiding principles for women offenders*. 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. & Austin, J. (1997).*Women in jail: Classification issues.* Washington, DC: USDOJ, National Institute of Corrections.

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

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.

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.

Van Dieten, Marilyn. 2008. *Women Offender Case Management Model.* Washington DC: National Institute of Corrections.

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 Iowa Department of Corrections. 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.