_____ **Research Report** _____

An Examination of a Reweighted
Custody Rating Scale for Women

An Examination of a Reweighted Custody Rating Scale for Women

Sara Rubenfeld

Correctional Service of Canada

February 2014

Acknowledgements

Executive Summary

Key words: *custody rating scale, validation, reweighting, women offenders, security level, custody classification*

The placement of offenders at correct security levels is a critical incarceration decision that influences inmates' living conditions, access to the community and programming options. To date, Correctional Service Canada has used the Custody Rating Scale (CRS), an actuarial tool, along with additional information to classify and place women into a security level at admission. However, this scale was originally developed with a sample of men and, because of this, critics argue against the use of this tool with women. Moreover, previous research demonstrates that a high number of Aboriginal women are recommended for placement in medium and maximum security levels. The goal of the current research is to examine the possibility of reweighting the CRS's items to improve its validity and relative efficacy for Aboriginal and non-Aboriginal women.

Women who had received an initial custody classification between June 2007 and April 2011 were randomly assigned to one of two groups, a construction sample and a validation sample. In total, 541 women (143 Aboriginal women and 398 non-Aboriginal women) were included in the construction sample and 542 women (143 Aboriginal women and 399 non-Aboriginal women) were included in the validation sample.

Reweighting of the CRS, using the Burgess method, was undertaken with the construction sample in order to give greater weight to those scale items that were most predictive of offenders' subsequent institutional adjustment. Following the reweighting, the validity of the scale was analyzed for both samples.

The results across both samples were consistent in demonstrating that reweighting increased the extent to which the CRS predicted subsequent institutional incidents and charges for the full sample of women and for the non-Aboriginal sub-sample. These improvements were not consistent for the Aboriginal sample. For example, reweighting had the effect of slightly reducing the disproportionate number of Aboriginal women at higher security levels. However, over-representation of Aboriginal women at medium and high levels of security continued to be present post-reweighting. Given these differing patterns and the limited magnitude of improvements, implementation of the reweighted scale is not recommended.

Alternative approaches to improving initial custody classification for women were discussed. Namely, the options of separately reweighting the scale for Aboriginal and non-Aboriginal samples or developing a new gender-informed instrument were raised as alternative methods that may improve the validity of an initial custody classification instrument.

Table of Contents

List of Tables

List of Figures

List of Appendices

Introduction

The security level assigned to an offender is a critical element of effective correctional management. Security classification levels influence inmates' living conditions (e.g., type of accommodation and freedom of movement within the institution), access to the community and programming options (Blanchette & Motiuk, 2004; Bonta & Motiuk, 1992; Hannah-Moffat & Shaw, 2001). Under- or over-classification of an offender could pose undue risks to the public, staff members and offenders or be inappropriately restrictive for offenders. Therefore, the placement of offenders at appropriate security levels is one of the most important decisions made after a person is admitted to prison.

Custody Rating Scale

In 1991, CSC implemented the Custody Rating Scale (CRS). This actuarial measure is designed to provide an objective, research-based instrument to aid in the determination of new federal offenders' security classification. In order to do so, the instrument is comprised of two dimensions: (1) an Institutional Adjustment (IA) sub-scale that consists of five items designed to assess risks associated with the offender committing institutional incidents and (2) a seven-item Security Risk (SR) sub-scale that measures the danger an offender would pose to the public if she escaped. This measure is completed, based on a review of an offenders' file, during the Offender Intake Assessment and yields a recommended security level. A caseworker will base a final classification recommendation on his or her clinical appraisal, available psychological assessments and the CRS result. Therefore, the security level that is designated by the CRS can either be approved or overridden by the case worker.[1]

As is often the case for risk-based measures (Hannah-Moffat & Shaw, 2001), the CRS was originally developed with a sample of men in federal penitentiaries (Porporino, Luciani, Motiuk, Johnston & Mainwaring, 1989). Validations of this instrument using a sample of men primarily have demonstrated significant convergence between factors that are measured differently (e.g., dynamic and static risk), but expected to be associated with offenders' CRS scores. As well, these assessments demonstrate increased occurrence of negative institutional

---

[1] The final placement decision is made by the warden or the Kikawinaw, the Institutional Head at the Okimaw Ohci Healing Lodge, based on the CRS recommendation and the case worker's clinical judgment.

outcomes, such as institutional incidents and escapes (Luciani, Motiuk & Nafekh, 1996[2]; see also Grant & Luciani, 1998); minor and major institutional incidents and serious institutional charges (Gobeil, 2011), at higher recommended security levels.

Applying the CRS to Women offenders

As the numbers of women incarcerated grows, the availability of larger sample sizes (i.e., greater numbers of women) provides sufficient statistical power to further validate the CRS and examine any deficiencies present when instruments oriented to men are applied to women.

Representation of Aboriginal Women

Previous examinations of the validity of the CRS for women have identified two main concerns: disproportionate representation of Aboriginal women at higher levels of security and weak associations between CRS recommendations and relevant outcomes.

In the first assessment that focused on the use of the CRS among women, Blanchette, Verbrugge and Wichmann (2002; see also Verbrugge & Blanchette, 2002) found that Aboriginal women scored significantly higher than non-Aboriginal women on both the IA and SR subscales. In turn, Aboriginal women were under-represented in minimum security designations and over-represented in medium and maximum security designations. Although this differential exists between Aboriginal and non-Aboriginal men, the extent of the discrepancy in security level recommendations is far greater for women (Barnum & Gobeil, in press; Gobeil, 2011).

These discrepancies between Aboriginal and non-Aboriginal women have led to some debate over whether Aboriginal women are over-classified or over-represented at higher security levels. *Over-representation* occurs when a greater proportion of one population is classified to a higher security level than another population. In contrast, *over-classification* refers to a situation where a population is placed at higher levels than what is necessary, given CCRA (1992) stipulations that classification levels must reflect the offender's institutional adjustment, escape risk and risk to the public in the event of an escape (Gobeil, 2008). Barnum & Gobeil (in press) argue that Aboriginal offenders are over-represented at higher security levels, rather than over-classified, because they have higher levels of criminogenic risk.

---

[2] Note that the sample size was limited to 65 women because very few women had a completed CRS at that time. The small size of this sample limits the extent to which conclusions can be drawn about the validity of the CRS for a female sample. Comparatively, 6,679 men were included in this validation study.

From another perspective, it is possible that over-classification may go undetected if high levels of criminogenic risk do not translate into lower levels of institutional adjustment, higher risk for escape and/or greater risk to the to the public in the event of an escape. Webster and Doob (2004) suggest that a discrepancy in security level assignment and placement would be justified if a greater proportion of Aboriginal women committed institutional incidents than non-Aboriginal women. However, the percentage of offenders involved in incidents, in Blanchette et al.'s 2002 study, was notably lower for Aboriginal offenders than non-Aboriginal offenders in both minimum (28.6% versus 40.0%, respectively) and medium (26.8% versus 52.4%, respectively) CRS designations.[3] Hence, although Aboriginal women's CRS scores and resulting security classifications were higher than were non-Aboriginal women's, fewer Aboriginal women were involved in institutional incidents, an outcome of which the CRS is intended to be predictive. This suggests that high levels of criminogenic risk that drive Aboriginal women into medium security may not translate into negative institutional adjustment.

Associations between CRS Recommendations and Relevant Outcomes

In examining the predictive validity of the CRS designations, Blanchette et al. (2002) found that, overall, the percentage of women involved in institutional incidents increased, as expected, from minimum to maximum security levels.[4] However, analysis of the IA and SR subscales and subscale items varied. In particular, the IA subscale was a significant predictor of violent and non-violent incidents for both ethnic groups. In contrast, the SR subscale was only a significant predictor of these outcomes for non-Aboriginal women. Further, the relationship between subscale items and violent and non-violent incidents varied such that the strength of the associations differed across ethnicity and were relatively weak. As well, the direction of some associations was counter-theory. In a follow-up investigation of the same cohort of women, Blanchette and Motiuk (2004) found significant, but deflated, linear patterns of increasing returns to custody, new convictions and new violent convictions, from minimum to maximum CRS recommendations.

More recently, Barnum and Gobeil (in press) found that the predictive ability of the CRS varied by the institutional outcome and by ethnicity. In some cases, the CRS was more predictive

---

[3] Although the reverse is true at maximum CRS designations, small sample sizes ($n_{aboriginal} = 6$, $n_{non-aboriginal} = 5$) limit the extent to which these results can be interpreted.

for Aboriginal women (e.g., predicting major institutional incidents) and in other cases, the CRS functioned better for non-Aboriginal women (e.g., predicting minor institutional charges). Results were more consistent when predicting institutional adjustment outcomes based on actual placement decisions.

Consistent with Blanchette et al.'s (2002) findings, Barnum and Gobeil (in press) stated that "the associations between individual items and outcomes of interest were weaker than expected" (p.38). Additionally, in some cases, the direction of the association was unexpected. A parallel examination of the CRS among men offenders (Gobeil, 2011), demonstrated that revalidations findings were stronger for men than women.

The Current Study

The results, outlined above, demonstrate that the CRS offers variable validity in the prediction of relevant outcomes for women offenders and inequivalence in the CRS between Aboriginal and non-Aboriginal women offenders. Concerns associated with this pattern of results, along with the continued use of the CRS in a form initially designed for men has prompted the Canadian Human Rights Commission (2003) to recommend the development of a security classification tool explicitly created for federally-sentenced women as well as actions that would address the disproportionate number of federally sentenced Aboriginal women classified as maximum security. To respond to these concerns, the goal of the current report is to reweight the CRS to improve its validity for Aboriginal and non-Aboriginal women.

Method

Participants

Participants included women who had received an initial custody classification decision between June, 2007 and April, 2011. The earlier time period was selected because it was the first month where the Offender Management System (OMS) provided a variable to link CRS scores and final security level decisions.

Thirty-five women with indeterminate sentences were excluded from this study because they did not have the necessary information to link CRS Scores and final security level decisions and because of poor data quality.

In total, 1,083 women offenders were included in the study, 286 were Aboriginal and 797 were non-Aboriginal. In order to divide the total sample into two sub-samples, a construction group and a validation group, stratified random sampling was used. This approach produced two randomly created sub-groups: the construction sample included 143 Aboriginal women and 398 non-Aboriginal women and the validation sample included 143 Aboriginal women and 399 non-Aboriginal women.

Data Source

The data used for this study were extracted from the offender records through the Correctional Service of Canada and Parole Board of Canada Data Warehouse. Data were obtained from the annual snapshot of the OMS for the 2010-2011 fiscal year. This snapshot of administrative data contains all computerized information pertinent to federal sentences served up to April 10, 2011. The data extracted contain information related to the demographics and incarceration characteristics of the women in the sample, security classification information, intake assessment results and institutional adjustment outcomes.

Measures

Demographic & Incarceration Characteristics

Several demographic and incarceration characteristics were examined in order to identify potential differences between the construction and validation samples.

Ethnicity. Offenders were classified into one of three groups: Caucasian, Aboriginal (Innu, Inuit, Métis and First Nations), and Other/Unknown (Arabic or Western Asian, Black, East Indian, Hispanic, Chinese, Filipino, Japanese, Korean, Latin American, South East Asian, Other and Unknown). Comparisons across ethnicity included two categories: Aboriginal, as define above, and non-Aboriginal (all other groups).

Age at sentence commencement. The age of offenders was calculated by subtracting the date of birth from the date of sentence commencement.

Region. The region of women offenders represents the region of the institution to which an offender was admitted. This variable is separated according to CSC's five regional divisions: Atlantic, Quebec, Ontario, Prairies and Pacific.

Marital status. Women were categorized into one of two groups: married (includes common-law) or single (includes divorced, separated and widowed). Women with an "unknown" marital status were removed from analyses pertaining to this variable.

Sentence length. The length of offenders' sentences was examined in two ways. First, a continuous measure of sentence length representing the number of years to be served for the current sentence was used. Those serving a life sentence have an indeterminate number of years to be served and are, therefore, not included in this variable. To accommodate this exclusion, a second categorical measure of sentence length that divides sentence length into one of three categories, three years or less, more than three years and life sentence, was used.

Offence type. Offenders' most serious offence type on their sentences was classified using three dichotomous variables: homicide, assault, and violent.

Homicide: offenders convicted of a homicide or related offences (i.e., homicide and related) were coded as "homicide" and other offence types were given a "non-homicide" code.

Assault: offenders convicted of an assault or related offences (i.e., sexual assault, major assault, and common assault) were coded as "assault" and other offence types were given a "non-assault" code.

Violent: offenders convicted of a violent offence or related offences (i.e., homicide and related, attempted murder, robbery, sexual assault, sexual abuse, kidnapping, abduction, weapon and explosives, major assault, common assault, and arson) were coded as "violent" and other offence types were given a "non-violent" code.

Security Classification

Items and variables related to security classification were incorporated in this study in order to perform the reweighting of the CRS and validation of the reweighted instrument.

Original Custody Rating Scale. The term "original CRS" is used here to represent the measure in its original form (i.e., prior to reweighting). This measure includes two subscales, an Institutional Adjustment (IA) subscale and a Security Risk (SR) subscale with five and seven items, respectively. The IA subscale is designed to predict risk of involvement in institutional incidents. The SR subscale is designed to assess the risk an offender would pose to the public if they escaped . The CRS is provided in Appendix A. An explanation of each item and its scoring can be found in CD 705-7 (2010).

Each subscale is scored separately by summing items found within the subscale. Distinct cut-off values are applied to each subscale to represent a recommended offender security level (OSL). When recommendations for the two subscales are inconsistent, the higher of the two values is retained.

Final security level. Three factors influence an offender's final security level placement decision: the original CRS, case workers' clinical judgement and the warden's discretion. Hence, the final decision about whether an offender should be initially placed in minimum, medium or maximum security can, and regularly does, differ from what is recommended by the CRS. It is recommended that the placement and recommendation not differ more than 20% of the time (Brennan & Austin, 1997).

Offender Intake Assessment

Measures of risk and reintegration potential were examined in order to assess the convergent validity of the security level indicated by the original CRS recommendation, the reweighted CRS recommendation and the final decision made. In all cases, where multiple risk and reintegration potential assessments were available, data from the assessment completed closest to the day of the CRS decision was examined.

Dynamic risk. Offenders' levels of criminogenic need are evaluated during intake in order to determine the level of intervention required. A rating of low, medium or high is given. It is expected that higher levels of dynamic risk will be associated with higher security level recommendations.

Static risk. A measurement of static factors, such as criminal history and severity of offence, is undertaken to capture offenders' levels of risk of re-offending. A rating of low, medium or high is given. Higher levels of static risk are expected to be associated with higher recommended security levels.

Institutional Adjustment Outcomes

Incidents and charges are two measures of offenders' behaviour while in an institution. As such, the association between the CRS and these outcomes speaks to the predictive validity of the actuarial measure.

Institutional incidents. Data on offenders who were perpetrators or associates in an institutional incident were captured. In this case, incident involvement included minor (e.g., theft, having unauthorized items, etc.), major incidents (e.g., inmate fights, hostage-taking, etc.) and other incidents (e.g., self-inflicted harm, security breaches or intelligence, etc.).

Institutional incidents were examined from two perspectives: (1) as a dichotomous variable that indicated whether the offender had at least one incident (coded as "1") or not (coded as "0") and (2) as a count variable that represented the number of incidents the offender had. It is expected that offenders with higher security level recommendations will be more likely to be involved in institutional incidents. As well, the number of days until an incident occurred was examined.

Institutional charges. Data from women convicted of a charge while incarcerated were included in the analysis. In this case, charges could be minor (e.g., disobeying roles, having an unauthorized item, disrespecting staff) or serious (e.g., intoxicant use, fights/assaults).[5] Institutional charges were examined from two perspectives: (1) as a dichotomous variable that indicated whether the offender had at least one charge (coded as "1") or not (coded as "0") and (2) as a count variable that represented the number of charges the offender had. It is expected that offenders with higher OSL recommendations will be more likely to have incurred institutional charges. As well, the number of days until a charge occurred was examined.

Analyses

After randomly dividing the full sample of participants into a construction and validation

---

[5] It should be noted that the choice between an institutional charge and/or incident is at correctional staff's discretion.

group, reweighting and validation of the construction group were undertaken. The first portion of this report begins with a thorough explanation of the Burgess reweighting method and concerns arising from the reweighting. Negative binomial regressions were used to determine the most suitable institutional adjustment outcome on which reweighting would be based. Once reweighting was completed, it was necessary to identify appropriate cut-points, delineating minimum, medium and maximum security. Cut-point determinations were made using Gallop's (2010) method for determining optimal cut-points along with a balancing of operational realities and consideration of high rates of false positives.

Validation analyses began with an examination of demographic and incarceration characteristics by ethnicity (Aboriginal or non-Aboriginal) and by OSL, as recommended by the reweighted CRS (minimum, medium or maximum). Following this, the reliability of the reweighted CRS was assessed using correlational analyses. Analysis of reliability is limited to an "eyeball" interpretation of correlations, given that measures of reliability (e.g., Cronbach's alpha) are biased when assessing scales with ordinal data (Zumbo, Gadermann & Zeisser, 2007).

Concordance rates between the original CRS, the reweighted CRS and final security level decisions were compared to assess concurrent validity. Convergent validity was investigated by exploring the association between the reweighted scale and measures of dynamic and static risk. Several analyses were conducted using institutional incidents and charges to examine the predictive validity of the reweighted CRS. Analysis of the concurrent, convergent and predictive validity were conducted for the overall sample and for Aboriginal and non-Aboriginal women separately.

In order to ensure that the results of the construction group were generalizable to a sample that was not used to reweight the scale, validation analyses were repeated among a second sample of offender (i.e., the validation group). These results were compared to the findings of the construction group.

In general, associations were assessed using Analysis of Variance (F tests), independent sample t-tests, Spearman rank correlations, Pearson chi-square (χ^2), Area under the Receiver Operating Curve analyses and survival analyses. In the case of chi-square, when 20% or more of the expected cell counts were less than 5, Fisher exact probabilities were used to assess significance rather than Pearson chi-square. When significant associations were found, Cramer's Phi (φ_c) and eta-squared (η^2) was used to assess the magnitude of the association between

9

variables. All analyses were conducted using SAS 9.2.

<center>**Results**</center>

Preliminary Results

    Reweighting the CRS

    The Burgess method (Nuffield, 1982) was employed to reweight the CRS. This method relies on providing a new weight to individual response options based on the extent to which they are more or less associated with an outcome of which the scale is intended to be predictive. In the past, this method has been used to reweight factors used to determine risk of recidivism (Nuffield, 1982), assess danger among "male intimate abusers" (i.e., men who had been violent toward their female partner; Lin, 2006) and develop a security reclassification scale for women (Blanchette & Taylor, 2005).

    Given that the CRS is intended to inform offender security levels, the outcome(s) of importance for the reweighting of the scale would be institutional adjustment, escape risk and risk to the public of a new offence if the offender were to escape. However, escapes occur so infrequently that instances of escape would not validly distinguish offenders at various security levels (Blanchette, 2004). Similarly, using "risk to the public for a new offence, if an escape were to occur" as an outcome of interest hinges on the assumption there would be a number of escapes and occurrences of reoffending during these escape periods. Therefore, the only viable outcome on which reweighting could be based is institutional adjustment.

    In past research (Blanchette, 2004; Gobeil, 2011; Grant & Luciani, 1998; Luciani et al., 1996), institutional adjustment has been measured by rates of institutional incidents and charges. Without a basis for selecting either outcome for reweighting, the first step in the current validation was to determine which outcome would be most appropriate for the reweighting of the CRS. To do this, the CRS was reweighted twice: based on charges and based on incidents (see Appendices B and C, respectively).

    Some psychometric concerns are apparent in the reweighted CRS. In particular, it would be expected that the reweighted values of scale items would increase in a linear fashion. For example, item 8 in Appendix B relates to the severity of the current offence. In the original CRS, offenders that have committed "minor or moderate" offences receive a value of 12, "serious or major" offences receive a value of 36 and "extreme" offences receive a value of 69. As scores increase, offenders are more likely to be placed in higher security levels. In turn, the reweighted

<center>11</center>

values increase linearly as severity of the current offence increases. That is, "minor or moderate" offences receive a score of -0.4, "serious or major" offences receive a score of 0.4 and "extreme" offences receive a score of 1.4. This indicates that there is consistent increase in the number of charges offenders receive while incarcerated, as the severity of their offences increase.

There are, however, several items that do not show a linear increase consistent with the CRS scoring. In some cases, this is due to a small number of women in a particular category. For example, the original CRS designates a value of 20 to "an escape or attempt from medium or maximum custody or an escape from minimum or community custody with actual or threatened violence over two years ago." However, there are no women in the current sample that meet this description. Similarly, in several other items, there are very small numbers of women meeting the various descriptions. This results in an unstable reweighting that may not be representative of the population.

Given this issue, some collapsing of response options was undertaken in order to deal with small sample sizes. In particular, three CRS items, "age at time of admission," "age at time of sentencing" and "prior parole and/or statutory release," have a large number of response options. Although having a sufficient sample size within each response option may have been feasible when creating and validating the measure with a sample of men, the relatively small number of women in federal institutions prohibits the use of these items in their original form. After examining the distribution of women across response options (see predictor distribution percentages in Appendices B and C) of the aforementioned items, their respective response options were collapsed into two or three response options. In general, this improves the linearity of the weighting and increases the number of women within each response options and, therefore, the stability of the weighting. However, collapsing across response options may reduce relevant distinctions that appear when more response options are used.

The most significant impact of collapsing of response options takes place in the "prior parole and/or statutory release" item when institutional incidents are used as the outcome of interest. For this item, response options were collapsed into two categories: (1) no prior parole and/or statutory release and (2) one or more prior parole and/or statutory release. In both cases, the Incident Rate per Predictor Category does not differ from the Base Incident Rate, which produces a weighting value of zero. As a result, the experience of having a prior parole and / or statutory release adds no predictive value to the reweighted CRS. As previously mentioned, this

may be due to the collapsing of response options minimizing relevant distinctions among response options. Alternatively, this result may indicate that having previously been on parole and/or statutory release does not predict women's involvement in institutional incidents.

Lastly, some items do not demonstrate linear progression of weightings, despite having a sufficient sample size to be confident about their representativeness of the population. For example, the reweighting of "number of prior convictions" based on charges (Appendix B; Item 6) increases linearly for the first four response options (i.e., none, one, two to four and five to nine). However, the final two response options (i.e., ten to fourteen and over fifteen) do not continue to show increased weightings. This may represent a form of a threshold that exists among women offenders such that the number of prior convictions is positively related to charges to a particular point; following this point, patterns of institutional charges are variable. In cases where there was a sufficient sample size across response options, yet a non-linear progression of weighting, no modifications were made to artificially create linearity in the weightings. This ensures that the weightings are driven by the extent to which they predict the institutional adjustment outcome, whenever sample size permits it.

Selecting an Outcome for Reweighting

Following the reweighting of the CRS, the two approaches (i.e., reweighting based on institutional charges and reweighting based on institutional incidents) were compared using three methods: (1) comparing the predictive value of individual scale items, by weighting approach, to their corresponding outcome, (2) contrasting the Bayesian Information Criterion (BIC; Raftery, 1995) obtained in pairwise comparisons of the subscale scores, obtained using both weighting approaches, in terms of their ability to predict institutional adjustment outcomes, and (3) examining ethnic differences in the experience of institutional incidents or charges, which may influence the subsequent validity of a reweighted CRS.

A Negative Binomial Regression[6] was conducted to assess individual scale items, post-weighting, in terms of their prediction of the total number of charges and total number of institutional incidents the offender obtained over the course of incarceration for a specific sentence. After acknowledging that the length of time incarcerated varies across offenders, a small selection of items significantly predicted the outcome on which the weighting was based

---

[6]Dispersion parameters obtained from the Negative Binomial regression demonstrated violations in the assumption of equidispersion, justifying the selection of a Negative Binomial regression in place of a Poisson regression.

(e.g., predicting total number of charges from items that are weighted based on charges). Namely, "history of involvement in institutional incidents," "alcohol / drug use" and "age (at the time of sentencing)," all items from the institutional adjustment subscale, were significant predictors of the number of charges offenders obtained over the course of incarceration for a particular sentence (Table 1). The Incidence Risk Ratio (IRR) for the first item, for example, indicates that for every one unit increase in "history of involvement in institutional incidents" the rate ratio for number of charges would be expected to decrease by a factor of 1.12 (12%).

Table 1

Regressing CRS Items Weighted for Charges on Total Number of Charges

Subscale	CRS Item (weighted based on charges)	Regression Coefficient (SE[a])	IRR[a]	95% Confidence Limits for IRR
Institutional Adjustment	History of involvement in institutional incidents	0.12 (0.03)	1.12****	1.07 – 1.19
	Escape History	0.03 (0.11)	1.03	0.83 – 1.29
	Street Stability	0.31 (0.17)	1.36	0.98 – 1.89
	Alcohol / drug use	0.26 (0.06)	1.29****	1.15 – 1.47
	Age (at the time of sentencing)	0.45 (0.15)	1.56*	1.17 – 2.13
	Number of prior convictions	0.15 (0.11)	1.16	0.94 – 1.44
	Most serious outstanding charge	0.14 (0.08)	1.15	0.99 – 1.33
	Severity of current offence	-0.19 (0.21)	0.82	0.55 – 1.25
	Sentence length	0.12 (0.14)	1.12	0.86 – 1.48
	Street stability	-0.35 (0.20)	0.70	0.48 – 1.04
Security Risk	Prior parole and / or statutory release	0.11 (0.32)	1.11	0.60 – 2.08
	Age (at time of admission)	0.06 (0.08)	1.06	0.90 – 1.25

[a] SE = standard error; IRR = Incidence Risk Ratio
*$p < .05$, **$p < .001$, ***$p < .0001$

For both the prediction of institutional charges (Table 1) and incidents (Table 2), the number of significantly predictive items (i.e., 3/12) was less than ideal. As a consequence, a small set of significant items garnered substantial weight and subsequent classification

15

recommendations were largely based on these few items rather than a range of variables thought to be important to the prediction of institutional incidents and charges. In an ideal situation, all or nearly all items being used to determine an offender's security level placement would predict relevant outcomes.

This issue aside, from the perspective of selecting one outcome over another, this analysis suggests that charges may be a more appropriate outcome measure than institutional incidents. Specifically, the issue associated with "prior parole and/or statutory release" for the reweighting based on institutional incidents, outlined earlier, reappears in this analysis in the form of zero scores (see Table 2). As well, the significant items in the reweighting based on institutional incidents are not as strongly associated with counts of institutional incidents as are the items in the reweighting based on charges in their prediction of the number of charges.

Table 2

Regressing CRS Items Weighted for Charges on Total Number of Incidents

Subscale	CRS Item (weighted based on charges)	Regression Coefficient (*SE*[a])	IRR[a]	95% Confidence Limits for IRR
Institutional Adjustment	History of involvement in institutional incidents	0.08 (0.03)	1.09*	1.03 – 1.15
	Escape History	0.12 (0.06)	1.13*	1.02 - 1.26
	Street Stability	-0.08 (0.17)	0.92	0.67 – 1.28
	Alcohol / drug use	0.11 (0.08)	1.12	0.96 – 1.30
	Age (at the time of sentencing)	0.36 (0.31)	1.44	0.78 – 2.66
	Number of prior convictions	0.16 (0.09)	1.18	0.99 – 1.40
	Most serious outstanding charge	0.22 (0.15)	1.24	0.93 – 1.67
	Severity of current offence	-0.70 (0.42)	0.49	0.22 – 1.12
Security Risk	Sentence length	0.04 (0.10)	1.04	0.85 – 1.27
	Street stability	0.38 (0.17)	1.47*	1.05 – 2.06
	Prior parole and / or statutory release	**0.00 (0.00)**	**0.00**	**0.00 – 0.00**
	Age (at time of admission)	0.08 (0.38)	1.08	0.51 – 2.28

[a] SE represents the standard error; IRR = Incidence Risk Ratio
*$p < .05$, **$p < .001$, ***$p < .0001$

In additional to examining the significance of regression coefficients associated with various reweighted items, the BIC values can also be used to guide model selection. When

analyzing BIC scores, the model with the smallest value is taken as the best fitting model (Coxe, West & Aiken, 2009).[7] In these analyses, the focus is on verifying whether the prediction of the two outcomes is strongest when using their respective approaches to weighting. For example, when predicting counts of incidents, the security risk subscale weighted based on institutional incidents (BIC = 1062.28) is lower than the security risk subscale weighted based on charges (BIC = 1080.52). In all cases, this expectation is met (see Table 3). Therefore, this analysis does not support or negate the selection of a particular approach over another.

Table 3

Predicting Institutional Adjustment Outcomes Using Both Weighting Approaches

	BIC[a] Scores for Total Number of Charges	BIC[a] Scores for Total Number of Incidents
Institutional Adjustment (weighted for charges)	1241.90	1058.69
Security Risk (weighted for charges)	1277.12	1080.52
Institutional Adjustment (weighted for incidents)	1254.45	1050.70
Security Risk (weighted for incidents)	1295.35	1062.68

[a] BIC = Bayesian Information Criterion

Lastly, from a less statistical perspective, there are ethnic differences in the experience of institutional incidents or charges. Specifically, while the percentage of Aboriginal and non-Aboriginal women experiencing charges is similar, fewer non-Aboriginal women have documented incidents than Aboriginal women (Table 4). Basing the weighting on an outcome, which does not occur as frequently for non-Aboriginal women, may influence the subsequent validity of a reweighted CRS. This finding would, therefore, support the use of charges over incidents.

---

[7] Note that the BIC tends to be biased against models with smaller sample sizes.

Table 4

Percentage of Negative Institutional Adjustment Outcomes by Ethnicity

	Aboriginal	**Non-Aboriginal**
Charges	34.27%	32.66%
Incidents	32.87%	25.13%

Taken together, the results of the first and third analyses lend support for the selection of charges as the institutional outcome of interest. Therefore, subsequent analyses will adopt the reweighting obtained by this measure (see Appendix B).

Determining Cut-Off Scores

Reviewing cut-points used to delineate offenders' security level is critical to reassessing the CRS. That is, reweighting of the CRS signifies a reconceptualization of the relative importance of items within the scale and the magnitude of the impact of various response options within the original CRS. In the same sense, reassessing cut-points represents a review of underlying assumptions about where appropriate cut-offs are in CRS scores for determining whether women should be placed in minimum, medium or maximum security. Also, reweighting produces final scores (i.e., weights) that differ substantially from the points allotted to response options within the original CRS. For example, the institutional adjustment subscale of the CRS varies between 0 and 186. After reweighting using the Burgess method, the institutional adjustment subscale produces final scores that vary between approximately -20 and 25. Therefore, use of original cut- points would not be practicable.

It was also necessary to review the cut-points because a decision was made to base security level recommendations on a total score, rather than on the larger of the two subscales. This decision was based on inconsistent findings of the predictive validity of the subscales, discussed earlier. There was particular concern about basing decisions on the SR sub-scale given that none of its items were significantly associated with the number of institutional charges received (Table 1). Also, the SR sub-scale often recommends placement in medium security over the minimum security recommendation provided by the IA sub-scale.

Various cut-points were examined by generating ROC curves using institutional incidents and charges to examine possible Optimal Operational Points (OOPs) for the reweighted CRS. In

an ideal situation, the reweighted measure would have the capacity to correctly identify offenders that will have institutional adjustment issues 100% of the time (i.e., true positives) and will never incorrectly label an offender as being likely to have institutional adjustment issues (i.e., false positives). In reality, actuarial measures are not 100% accurate in their ability to predict outcomes and, therefore, cost-benefit considerations must be taken into account.

In the case of security classifications, serious consequences can come from incorrect decisions (i.e., false positives and false negatives). Therefore, sensitivity[8] and specificity[9] constraints were also considered in the process of selecting cut-off scores in order to balance efforts to place women in an appropriate incarceration environment with the safety of staff and offenders within the institution.

Gallop (2010) presents a method for determining OOPs that simultaneously maximizes both sensitivity and specificity. In this situation, the consequences of false positives and false negatives are considered to be approximately equal. Therefore, maximizing sensitivity and specificity would simply entail identifying the point at which the average of these two values is at its highest. For example, when using this approach to determine an appropriate cut-point to distinguish placement in medium versus maximum security, Appendix D demonstrates that the mean of sensitivity and specificity is maximized when the reweighted CRS score is 0.8. This cut-point would result in approximately 40.5% of all women being placed in maximum security, a substantial shift from conventional practice.

Given that a cut-off determination based on maximizing sensitivity and specificity results in unacceptably high percentages of women being placed in maximum security, alternative approaches were considered. Gallop presents an approach where a pre-assigned value for sensitivity or specificity is used, based on the consequences associated with false negatives and false positives. Prioritizing false negatives would entail setting a high minimum sensitivity score in order to correctly identify offenders who have institutional adjustment events. In contrast, prioritizing false positives would emphasize setting a high minimum specificity score in order to ensure that offenders who do not have institutional adjustment events are not over-classified.

The decision about prioritization was largely based on concerns associated with the

---

[8] Sensitivity refers to the proportion of correctly identified true positives (e.g., correctly indentifying offenders who will have an institutional charge).

[9] Specificity refers to the proportion of correctly identified true negatives (e.g., correctly identifying offenders who will not have an institutional charge).

validity of the reweighted CRS. These concerns, discussed earlier in the context of reweighting, resurface when examining the Receiver Operating Curve (ROC) information. As demonstrated in the example in Appendix D, the probability level of a charge within three months only exceeds .60 for offenders with the highest obtained reweighted CRS score. On a more general level, these probability levels indicate that even offenders with high CRS scores are only somewhat likely to display negative institutional adjustment outcomes. This may reflect the low significance of reweighted CRS items, discussed earlier, or it may also indicate that the use of charges as a measure of institutional adjustment has limited relevance for women offenders.[10]

The low probability of obtaining a charge within three months from the decision date also manifests itself in true and false positives. Whereas one would hope to see that the number of correctly predicted events would far exceed the number of non-events predicted as events, this is not the case with the reweighted CRS scores. Across many reweighted CRS scores, the number of false positives matches exceeds the number of true positives. Therefore, the risk of making an error increases at a steeper rate than does the number of correct predictions (see Figure 1).

Figure 1. Visual Representation of False Positive and False Negative Progression across Reweighted CRS Score
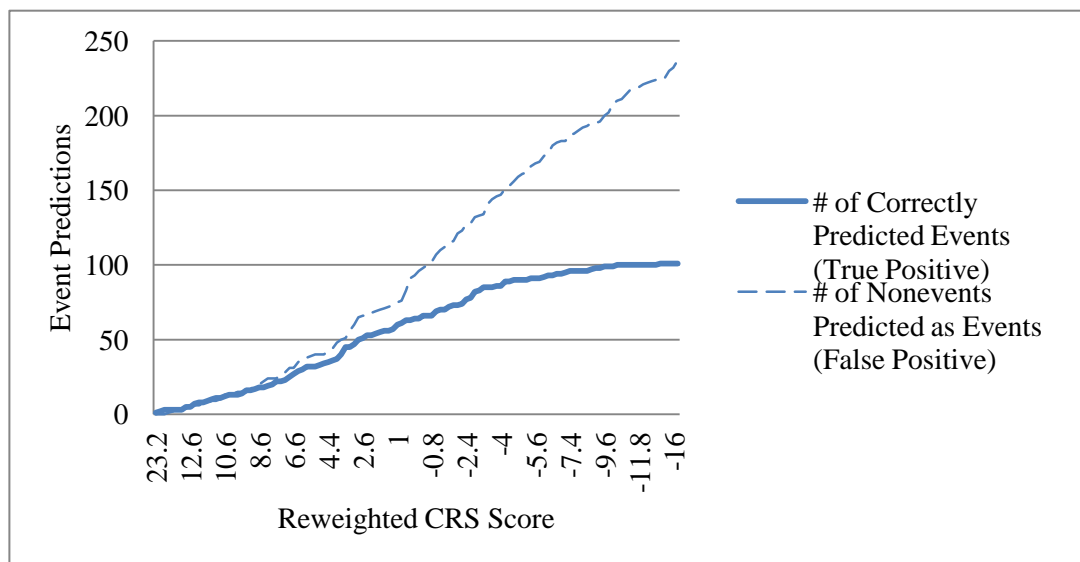
Based on these concerns, priority was given to false positives to ensure a more

---

[10] Note that the ROC data associated with institutional incidents show weaker patterns of low probability levels (e.g., the highest reweighted CRS scale score obtained (23.2) is associated with a 0.54 probability level).

conservative application of the measure (e.g., to minimize over-classification). At the same time, current operational procedures were considered as major deviations from current distributions of offenders across security levels (see "final decisions" in Table 7) could pose unnecessary risks to offenders and staff. Table 5 summarizes recommended scale scores for the three security levels. Sensitivity and specificity values for the two cut-points (i.e., delineating minimum from medium and medium from maximum) are also provided.

Table 5

Recommended Cut-Points for Weighted Scores

Recommended Security Level	Weighted Score	Sensitivity	Specificity
Minimum	Below 0		
Medium	From 0 to less than 11.6	0.70	0.60
Maximum	11.6 and greater	0.14	0.96

Rescaling the Scale

Given that the reweighted CRS produces scores with a range that is vastly different from the original CRS, rescaling was applied such that a value of 17 was added to each score. The new sum was then multiplied by 5 in order to eliminate decimal points. Table 6 summarizes the updated cut-points based on the rescaled score.

Table 6

Cut-Points Based on the Rescaled Score

Recommended Security Level	Rescaled Score
Minimum	below 85
Medium	from 85 to less than 143
Maximum	143 and greater

Using these cut-off scores, the reweighted CRS produces distributions that are similar to the original CRS recommendation and final OSL decision (Table 7). However, slightly greater percentages of offenders tend to be placed in minimum and maximum security using the reweighted scale than what is found in the other two distributions. In turn, a slightly lower

percentage of offenders would be recommended to medium security using the reweighted scale than the original CRS recommendation and the final OSL decision.

Table 7

OSL Distributions Based on the Reweighted CRS, Original CRS and Final Decisions

	Reweighted CRS % (*n*)	**Original CRS % (*n*)**	**Final Decision % (*n*)**
Minimum	49.17 (266)	48.98 (265)	46.58 (252)
Medium	44.92 (243)	46.21 (250)	48.98 (265)
Maximum	5.91 (32)	4.81 (26)	4.44 (24)

n = sample size

Upon completion the reweighting, cut-off determination and rescaling, subsequent analyses were conducted to assess the validity of the reweighted scale. First, the same sample which was used to develop the reweighted measure (construction sample) will undergo validation testing. Next, to ensure that the finding of the construction sample are generalizable to a distinct sample of women offenders, a second (validation sample) sample, which had no influence over the reweighting of the CRS, will undergo equivalent validation analyses.

Main Analyses: Construction Sample

Descriptive/Offence Information for Construction Sample

This section examined demographic and incarceration characteristics in order to identify any unique characteristics among Aboriginal and non-Aboriginal women offenders (Table 8). The majority of women (58.6%) were Caucasian, followed by Aboriginal (26.4%) and other (e.g., black, Southeast Asian) or unknown (15.0%). The Aboriginal women in the current sample tended to be younger at the time of sentence commencement than the non-Aboriginal women, admitted into an institution in the Prairie Region and are more likely to have crimes that involve homicide, violence and/or assault than non-Aboriginal women.

Table 8

Demographic Characteristics by Ethnicity

	Aboriginal *% (n)* or *M (SD)*	Non-Aboriginal *% (n)* or *M (SD)*	Association Between Variables
Demographic Characteristics			
Ethnicity			
Aboriginal	26.43 (143)	-	
Caucasian	-	58.60 (317)	
Other or Unknown	-	14.97 (81)	
Age at Sentence Commencement	30.31 (8.37)	34.48 (9.91)	$t_{(294)} = -4.86^{***}$, $\eta^2 = 0.07$
Marital Status			
Has Partner	33.80 (48)	39.65 (157)	$\chi^2_{(2)} = 1.51$
Single	66.20 (94)	60.35 (239)	
Region			
Atlantic	4.90 (7)	19.60 (78)	
Quebec	4.90 (7)	25.38 (101)	$\chi^2_{(4)} = 110.85^{***}$, $\Phi c = 0.45$
Ontario	14.69 (21)	26.88 (107)	
Prairies	61.54 (88)	18.34 (73)	
Pacific	13.99 (20)	9.80 (39)	
Incarceration Characteristics			
Length of Aggregate Sentence (Years)[a]	3.04 (1.9)	2.92 (1.39)	$t_{(196)} = 0.68$
Length of Aggregate Sentence			
3 Years or Less	62.24 (89)	65.40 (259)	
More Than 3 Years	37.76 (54)	34.09 (135)	$p = 0.72$[b]
Life Sentence	0.00 (0)	0.51 (2)	
Type of Offence			
Assault	24.48 (35)	10.41 (41)	$\chi^2_{(2)} = 17.09^{***}$, $\Phi c = 0.18$
Non-Assault	75.52 (108)	89.59 (353)	
Homicide	12.59 (18)	3.81 (15)	$\chi^2_{(2)} = 14.02^{***}$, $\Phi c = 0.16$
Non-Homicide	87.41 (125)	96.19 (379)	
Violent	62.94 (90)	32.49 (128)	$\chi^2_{(2)} = 40.34^{***}$, $\Phi c = 0.27$
Non-Violent	37.06 (53)	67.51 (266)	

M = mean, SD = standard deviation, n = sample size; Column totals may not add to 100% due to rounding. [a] Excludes life sentences. [b] This p-value represents the significance associated with the Fisher's exact test. * $p < 0.05$, ** $p < 0.001$, *** $p < 0.0001$

In addition to examining ethnic differences in demographic and incarceration characteristics, it is important to investigate whether there are appreciable differences between

offenders in minimum, medium and maximum security levels. In part, this information provides insight into the characteristics of offenders across security levels. At the same time, many of the demographic variables examined here are included in or influence the CRS. Therefore, it would be expected that patterns of differences across security level would mimic patterns of scoring of the original and reweighted CRS.

The original CRS asserts a negative relationship between institutional adjustment and age such that women are allotted more points if they are admitted at a younger age. Consistent with this assumption, when using the reweighted CRS recommendation, the average age of women is highest at minimum security ($M = 36.58$ years), followed by medium security ($M = 30.65$ years) and maximum security ($M = 27.44$ years; see Table 9). Post-hoc analyses, using the Tukey test, indicate that significant results (F (2,538) $= 33.83$, $p < .0001$, $\eta^2=0.11$) are explained by differences between women in minimum and medium security and minimum and maximum security.

Results provided in Table 9 demonstrate patterns that are consistent with the underlying empirically driven assumptions of the CRS; namely that longer sentences and more serious offences are associated with being placed in higher security levels. Although length of aggregate sentence (in years) does not differ significantly across OSL, there is a pattern of increasing sentence length across minimum ($M = 2.88$), medium ($M = 2.94$) and maximum ($M = 3.65$) security. The non-significance of these results may, in part, be due to the distribution of women's sentence lengths. Positive skewness and kurtosis of this variable indicates that sentences tend to cluster at shorter lengths, which would produce low average (mean) sentence lengths.

There are significant differences in the percentage of women with offences that involve violence, assault and/or homicide across OSLs, as recommended by the reweighted CRS. Lastly, there is a significant association (χ^2 (8, $n = 541$) $= 25.66$, $p < .001$, $\Phi c = 0.15$) between the reweighted tool's security level recommendations and region such that greater proportion of women offenders in the Prairie region are recommended to be placed in medium and maximum security than in other regions. This is likely related to the higher proportion of Aboriginal offenders, who have generally committed more serious crimes, being located in the Prairie region. In contrast, the region of Ontario has the highest proportion of women placed in minimum security.

Table 9[11]

Characteristics of Offenders across OSL Recommendations, Based on the Reweighted CRS

	Reweighted CRS Recommendation			Association Between Variables
	Minimum	**Medium**	**Maximum**	
	% (n) or M (SD)	% (n) or M (SD)	% (n) or M (SD)	
Demographic Characteristics				
Age at Admission	36.58 (10.55)	30.65 (7.75)	27.44 (5.99)	$F_{(2,538)} = 33.83$***, $\eta^2 = 0.11$
Region				
Atlantic	13.91 (37)	17.28 (42)	18.75 (6)	
Quebec	19.17 (51)	20.99 (61)	18.75 (6)	$\chi^2_{(8)} = 25.66$**, $\Phi c = 0.15$
Ontario	32.71 (87)	14.81 (36)	15.63 (5)	
Prairies	24.44 (65)	35.39 (86)	31.25 (10)	
Pacific	9.77 (26)	11.52 (28)	15.63 (5)	
Incarceration Characteristics				
Length of Aggregate Sentence (Years)[bc]	2.88 (1.26)	2.94 (1.55)	3.65 (3.08)	$F_{(2,536)} = 2.20$
Length of Aggregate Sentence				
3 Years or Less	50.00 (16)	66.12 (160)	64.91 (172)	
More Than 3 Years	50.00 (16)	33.06 (80)	35.09 (93)	$p = 0.18$[d]
Life Sentence	0.00 (0)	0.83 (2)	0.00 (0)	
Offence Type				
Violent and Related	25.86 (68)	50.83 (123)	84.38 (27)	$\chi^2_{(2)} = 59.63$***, $\Phi c = 0.33$
Non-Violent	74.14 (195)	49.17 (119)	15.63 (5)	
Assault and Related	7.60 (20)	17.77 (43)	40.63 (13)	$\chi^2_{(2)} = 30.34$***, $\Phi c = 0.24$
Non-Assault	92.40 (243)	82.23 (199)	59.38 (19)	
Homicide and Related	5.32 (14)	6.61 (16)	9.38 (3)	$\chi^2_{(2)} = 0.98$
Non-Homicide	94.68 (249)	93.39 (226)	90.63 (29)	

M = Mean; *SD* = standard deviation. Column totals may not add to 100% due to rounding. [b] Excludes life sentences. [c] This variable was logarithmically transformed due to violations of the assumption of normality. The means and standard deviations presents were untransformed. [d] This p-value represents the significance associated with the Fisher's exact test.
*$p < 0.05$, **$p < 0.001$, ***$p < 0.0001$

---

[11] Note that both *age at admission* and *aggregate sentence length* were non-normal. Age at admission was positively skewed and aggregate sentence length was both positively skewed and kurtotic. A square root transformation produced normal distribution for age at admission, but a logarithmic transformation was required to achieve a near normal distribution for aggregate sentence length. Parametric tests (i.e., independent sample t-test and analysis of variance) were applied to the transformed variables. However, with the exception of one test of significance, these tests did not produce results that differed substantively from the original (untransformed) data. In examining group differences in aggregate sentence length across security levels, results that were significant using the non-normal data became insignificant after logarithmically transforming this variable. In this case, the transformed variable was retained. In all other cases, the untransformed variables were retained.

Reliability of the Reweighted CRS

Since its conception, the CRS has been treated as a bidimensional measure, with an institutional adjustment subscale and security risk subscale intended to assess risks associated with these two respective constructs. Typically, a measure of reliability, such as Cronbach's alpha, would assess the internal consistency of items within a scale or subscale. However, Cronbach's alpha tends to be downwardly biased when assessing scales with ordinal scaling, particularly with items that have few response options (Zumbo et al., 2007). In lieu of such a measure, inter-item Spearman rank correlations are examined (see Appendix E).

The correlation among items varies widely, from -0.14 to + 0.93. Slightly over one-third (36%) of all inter-item correlations are significant. Certain items, such as incident involvement, prior parole and/or statutory release and number of prior convictions are significantly associated with many scale items. However, other items (e.g., most serious outstanding charge, sentence length, age (at time of admission) and age (at time of sentencing) show little to no association with other scale items. In these cases, this would suggest poor coherence within the components of the scale.

In Appendix E, items belonging to the same sub-scale are demarcated using a thicker border in order to draw attention to items that are intended to measure the same underlying constructs. It would be expected that items that tap into the same construct would show some level of association, without being singular (i.e., redundant) in their focus. In both cases, the associations between the majority of items within these groupings are non-significant. Although this analysis is only preliminary in nature, it suggests that items may not consistently measure their intended construct. This finding supports the use of a total score to delineate security level recommendations, rather than two sub-scales. Further examination of the factor structure of these items would shed light on which items best support the measurement of institutional adjustment and security risk.

Classification across Ethnicity

Reweighting the CRS offers the opportunity to examine the question about whether Aboriginal women are over-classified or over-represented from another perspective. Specifically, it is possible that over-classification may go undetected if high CRS scores do not translate into lower levels of institutional adjustment, higher risk for escape and/or greater risk to the public in

the event of an escape. Placing greater weighting on items that are relevant to the prediction of institutional adjustment creates a scale that is more predictive of institutional adjustment. Hence, recommendations stemming from this reweighted measure should coincide more closely with appropriate classification levels than the original CRS.

Table 10 demonstrates that although the association between OSL and ethnicity is significant, with a small to medium effect size, in all three cases, the degree of association is lower when basing recommendations on the reweighted measure than when basing recommendations or actual placements on the original measure. The fact that the reweighted scale produces a weaker association between OSL and ethnicity than the other two measures that are based on or influenced by the original CRS suggests that some over-classification may have been occurring. However, even when efforts were made to improve the linkage between the CRS and institutional adjustment, the distribution of Aboriginal women at higher security levels is still greater than the distribution of non-Aboriginal women. This suggests that much of the classification of Aboriginal women at higher security levels is driven by over-representation.

Table 10

Associations between Security Level Classifications and Ethnicity

		Aboriginal (%)	Non-Aboriginal (%)	χ^2 (*df*)	φ_c
Based on the Reweighted CRS	Minimum	27.27	57.04		
	Medium	64.34	37.94	37.29 (2)***	0.26
	Maximum	8.39	5.03		
Based on the Original CRS	Minimum	25.17	57.54		
	Medium	67.83	38.44	44.09 (2)***	0.29
	Maximum	6.99	4.02		
Final Decision	Minimum	21.68	55.53		
	Medium	70.63	41.21	49.12 (2)***	0.30
	Maximum	7.69	3.27		

df = degrees of freedom
*$p < 0.05$, **$p < 0.001$, ***$p < 0.0001$

Concordance Rates (Concurrent Validity)

Past validations of the CRS have examined concurrent validity by comparing CRS recommendations and final decisions (i.e., initial OSL placements). In these cases, the occurrence of overrides of the CRS recommendation would indicate that the use of clinical judgment and/or the warden's discretion resulted in a decision that disagreed with the measure's recommendation. In the current context, concordance can be looked at from a couple of perspectives. First, if the concordance rates between the reweighted scale and final decisions were higher than the rates between the original scale and final decisions, it would suggest that there may be value added by the reweighting such that its recommendations give greater weight to those aspects that are important to making final OSL determinations.

The rates obtained (see Table 11) do not demonstrate greater concordance between the reweighted scale and the final decision. Concordance between the original CRS and the final decision (overall percentage = 73.75) were higher than that which was found between the reweighted CRS and final decision (overall percentage = 70.97). It is difficult to determine whether this lower concordance is due to the quality of the reweighted scale or the fact that final decisions incorporated recommendations of the original CRS and not the recommendations of the reweighted CRS.

From a separate perspective, a comparison of concordance rates indicates how reweighting influences recommendations across ethnicity (also see Appendix F for further detail). When it comes to Aboriginal offenders, the reweighted CRS was more likely to recommend placement in a lower level (lower classification = 12.59%) than higher OSL levels (higher classification = 11.89%) compared to the original CRS recommendation. A similar result was found for Aboriginal offenders when comparing placement in lower security levels (18.89%) to placement in higher security levels (14.68%) between the reweighted CRS and final decisions. In contrast, a greater percentage of non-Aboriginal offenders were classified to higher levels, when comparing the reweighted CRS to the original CRS (15.07%) and the final decision (13.82%), than classified to lower levels (13.57% for both the original CRS and final decision). These results echo earlier findings (Table 11), which demonstrated that a scale that was reweighted to increase its predictiveness of institutional adjustment recommended that Aboriginal women be placed in lower security levels than did the original CRS and the final decision. In contrast, a greater proportion of non-Aboriginal offenders were recommended to

higher security levels.

Table 11

Percentage of Concordance, Lower Classification and Higher Classification

		Concordance %	Lower Classification %	Higher Classification %
Between Original CRS and Reweighted CRS	Overall Percentages	72.46	13.31	14.24
	Aboriginal Percentages	75.52	12.59	11.89
	Non-Aboriginal Percentages	71.35	13.57	15.07
Between Final Decision and Reweighted CRS	Overall Percentages	70.97	14.97	14.05
	Aboriginal Percentages	66.44	18.89	14.69
	Non-Aboriginal Percentages	72.61	13.57	13.82
Between Original CRS and Final Decision	Overall Percentages	73.75	14.05	12.19
	Aboriginal Percentages	72.73	15.39	11.89
	Non-Aboriginal Percentages	74.13	13.57	12.31

Convergent Validity

The association between the reweighted CRS and measures of dynamic and static risk were examined to assess convergent validity (Table 12). As would be expected, in all cases, there was a significant association between these variables, both overall and across ethnicity.

Specifically, static and dynamic risk increased when comparing offenders in minimum security to medium security and medium security to maximum security. However, a large percentage of offenders are recommended for placement in minimum security, despite demonstrating medium to high levels of dynamic and static risk. This is particularly true for dynamic risk levels among Aboriginal offenders; over 90% of this group obtained medium to high dynamic risk scores, but would be recommended to minimum security.

This finding may be related to the earlier proposition that high levels of dynamic and static risk may not translate into institutional incidents and charges. If they do not, items in the CRS that would capture aspects of dynamic and static risk would receive lower weightings and could produce the findings seen here.

Table 12

Associations between Security Levels (Recommendations and Placement) and Measures of Risk and Reintegration Potential

		Dynamic Risk			Static Risk		
		Low	Medium	High	Low	Medium	High
		n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Reweighted CRS: Overall	Minimum	25.21 (61)	49.59 (120)	25.21 (61)	53.99 (142)	34.22 (90)	11.79 (31)
	Medium	1.84 (4)	33.64 (73)	64.52 (140)	20.42 (49)	56.25 (135)	23.33 (56)
	Maximum	0 (0)	22.22 (6)	77.78 (21)	6.45 (2)	35.48 (11)	58.06 (18)
χ^2 (df)			105. 22 (4)***			94.70 (4)***	
φ_c			0.33			0.30	
Reweighted CRS: Aboriginal	Minimum	9.68 (3)	41.94 (13)	48.39 (15)	48.72 (19)	33.33 (13)	17.95 (7)
	Medium	1.25 (1)	22.50 (18)	76.25 (61)	13.19 (12)	50.55 (46)	36.26 (33)
	Maximum	0 (0)	36.36 (4)	63.64 (7)	16.67 (2)	25.00 (3)	58.33 (7)
Fisher's Exact		p = 0.03*			p = 0.0002***		
Reweighted CRS: Non-Aboriginal	Minimum	27.49 (58)	50.71 (107)	21.80 (46)	54.91 (123)	34.38 (77)	10.71 (24)
	Medium	2.19 (3)	40.15 (55)	57.66 (79)	24.83 (37)	59.73 (89)	15.44 (23)
	Maximum	0 (0)	12.50 (2)	87.50 (14)	0 (0)	42.11 (8)	57.89 (11)
χ^2 (df)			79.16 (4)***			67.38 (4)***	
φ_c			0.33			0.29	

df = degrees of freedom, n = sample size
*$p < 0.05$ **$p < 0.001$ ***$p < 0.0001$

Predictive Validity

The predictive ability of the reweighted CRS was assessed using three methods, across both outcomes variables of interest: institutional incidents and charges. First, chi-square tests of independence analyses on the full sample were used to examine associations between the occurrence of negative institutional adjustment outcomes (i.e., incidents or charges), at any point during the incarceration period, and CRS recommended security levels. Second, Area Under the ROC Curves (AUCs) were used to examine the CRS's capacity to discriminate among the offenders' likelihood of having incidents and charges that occur within three months of their placement in the security level.[12] Third, survival analyses allowed for a determination of whether the amount of time to an institutional incident or charge depended on the CRS recommendation. In all three sets of analyses, recommendations based on both the original and reweighted CRS were contrasted to determine whether predictive ability was enhanced by the reweighting of the CRS.

Institutional Incidents

Patterns of significant associations between institutional incident involvement and CRS recommendations were similar across the reweighted and original scale (see Table 13). In both cases, there was a significant association for non-Aboriginal women and for women overall. However, there was an increase in the magnitude of the significant associations, using the reweighted CRS. Namely, the effect size increased from 0.13 and 0.14 when using the original CRS to 0.19 and 0.21 when using the reweighted CRS, overall and for non-Aboriginal offenders, respectively. Put otherwise, the distribution of offenders with institutional incidents becomes more distinctive across recommended security levels when the reweighted scale is applied.

In contrast, results were not significant for Aboriginal women. Rather, as found in previous research (Hardyman, 2001), similar levels of involvement in institutional incidents were demonstrated across minimum and medium security levels.

---

[12] "Within three months of their OSL decision" refers to within three months from the latter of the date the OSL decision was made or the date of admission.

Table 13

Association between the CRS (Original and Reweighted) and Institutional Incidents

	Cases Involving Institutional Incidents %		
	Overall	**Aboriginal**	**Non-Aboriginal**
Reweighted CRS			
Minimum	18.80 (50/266)	28.21 (11/39)	17.18 (39/227)
Medium	34.57 (84/243)	33.7 (31/92)	35.10 (53/151)
Maximum	40.63 (13/32)	41.67 (5/12)	40.00 (8/20)
χ^2 (*df*)	19.07 (2)***	0.83 (2)	17.95 (2)***
φ_c	0.19		0.21
Original CRS			
Minimum	21.89 (58/265)	30.56 (11/36)	20.52 (47/229)
Medium	31.20 (78/250)	32.99 (32/97)	30.07 (46/153)
Maximum	42.31 (11/26)	40.00 (4/10)	43.75 (7/16)
χ^2 (*df*)	8.80 (2)*	0.32 (2)	7.51 (2)*
φ_c	0.13		0.14

df = degrees of freedom
*$p < 0.05$ **$p < 0.001$ ***$p < 0.0001$

A three month period was selected for examining AUCs in order to provide a sufficient period of time for incidents to occur while avoiding the loss of participants who were no longer being observed.[13] In total, 62.48% (338/541) of offenders were observed for three months following their OSL decision. By ethnicity, 60.14% (86/143) and 63.32% (252/398) of Aboriginal and non-Aboriginal offenders were observed for three months following their OSL decision, respectively.

The AUC can range from 0.5 (random discrimination) to 1.0 (perfect discrimination; Swets, 1988). In practice, a measure that produces an AUC of 0.6 or greater is considered an 'acceptable' predictor and a measure with an AUC of 0.7 or greater is considered a 'good'

---

[13] In the current study, the end of observation would occur either because the offender has been released, has received a new security level assessment or remained in the institution following the end of the data collection period, April 10, 2011.

predictor. Values presented in Table 14 demonstrate that the reweighting improved the prediction of incident occurrence. In line with the previous chi-square analyses, the reweighted scale produced acceptable AUCs values for the group, as a whole, and for non-Aboriginal women. Although the AUCs increased from below acceptable, with the original CRS, to acceptable, with the reweighted CRS, a non-significant chi-square for the ROC contrast estimation (χ^2 (2, n = 541) = 1.75, $p > 0.05$) indicates that there was not a significant difference between these two measures.[14]

Table 14

AUCs for the Prediction of Institutional Incidents using the CRS (Original and Reweighted)

	Overall	Aboriginal	Non-Aboriginal
Reweighted CRS: AUC (95% CI)	0.61 (0.54-0.68)	0.58 (0.46-0.70)	0.61 (0.53-0.70)
Original CRS: AUC (95% CI)	0.54 (0.46-0.61)	0.55 (0.43-0.67)	0.55 (0.46-0.63)

AUC = Area under the ROC Curve, CI = Confidence Interval

Survival analyses produced similar results as above. The original CRS and the reweighted CRS were significantly associated with time in both the overall and the non-Aboriginal sample (Table 15). In both cases, however, the level of significance increased with the reweighted scale.

---

[14] Results for non-Aboriginal (χ^2 (2, $n = 398$) = 0.83, $p > 0.05$) and Aboriginal (χ^2 (2, $n = 143$) = 0.08, $p > 0.05$) women were also not significant.

Table 15

Survival Time to Incident, Overall and by Ethnicity, using the CRS (Original and Reweighted)

	N	% failed (*n*)	Test of Equality χ^2 (*df*)	Uncensored *M* (*SE*)	Censored *M* (*SE*)
Overall					
Reweighted					
Minimum	266	18.80 (50)		101.60 (15.00)	144.80 (7.47)
Medium	243	34.57 (84)	26.11 (2)***	104.50 (99.07)	118.70 (8.94)
Maximum	32	40.63 (13)		69.61 (25.62)	62.47 (19.00)
Original					
Minimum	265	21.89 (58)		86.36 (13.01)	137.60 (7.57)
Medium	250	31.20 (78)	9.35 (2)*	113.60 (11.55)	127.80 (8.77)
Maximum	26	42.31 (11)		81.36 (28.97)	57.93 (20.14)
Aboriginal					
Reweighted					
Minimum	39	28.21 (11)		109.60 (27.72)	129.70 (16.13)
Medium	92	33.70 (31)	4.25 (2)	97.74 (17.62)	104.80 (14.15)
Maximum	12	41.67 (5)		58.00 (23.51)	13.86 (45.84)
Original					
Minimum	36	30.56 (11)		82.45 (25.67)	114.70 (19.11)
Medium	97	32.99 (32)	1.45 (2)	104.40 (17.54)	110.60 (13.24)
Maximum	10	40.00 (4)		69.25 (27.00)	45.67 (40.50)
Non-Aboriginal					
Reweighted					
Minimum	227	17.12 (39)		99.34 (17.73)	147.00 (8.25)
Medium	151	35.10 (53)	19.28 (2)***	108.50 (13.78)	127.40 (11.49)
Maximum	20	40.00 (8)		76.88 (40.13)	73.17 (22.94)
Original					
Minimum	229	20.52 (47)		87.28 (15.01)	140.80 (8.19)
Medium	153	30.07 (46)	7.30 (2)*	120.00 (15.40)	138.20 (11.50)
Maximum	16	43.75 (7)		88.29 (44.36)	66.11 (21.87)

df = degrees of freedom, M = Mean, SE = Standard Error, n = sample size, uncensored M = the average number of day to have an event, censored M = average number of days at risk before period ended

*$p < 0.05$ **$p < 0.001$ ***$p < 0.0001$

The results for the overall sample and non-Aboriginal sub-sample demonstrated similar patterns because non-Aboriginal women account for the majority of the overall sample. As a result, the sub-sample of Aboriginal women has a weaker influence on overall patterns and are, therefore, not well-represented by overall results. The three sets of analyses pertaining to the predictive validity of the CRS (original and reweighted) for Aboriginal women were similar in their non-significant findings. Although there was some movement toward significant values (e.g., χ^2 (2, n= 143) = 1.45 for the original CRS and χ^2 (2, n = 143) = 4.25 for the reweighted CRS, Table 15), the reweighted scale did not go far enough in improving the CRS to differentiate incident occurrences among Aboriginal women across security levels.

Institutional Charges

The results of the three sets of analyses pertaining to institutional charges were similar to findings as they relate to institutional incidents. Namely, the overall sample and non-Aboriginal sub-sample are significantly associated with OSL recommendations given by the original and reweighted CRS (see Table 16). As well, these associations were stronger with the reweighted scale than the original scale.

Table 16

Association between the CRS (Original and Reweighted) and Institutional Charges

	Cases Involved in Institutional Charges %		
	Overall	**Aboriginal**	**Non-Aboriginal**
Reweighted CRS			
Minimum	22.56 (60/266)	28.21 (11/39)	21.59 (49/227)
Medium	42.39 (103/243)	35.87 (33/92)	46.36 (70/151)
Maximum	50.00 (16/32)	41.67 (5/12)	55.00 (11/20)
χ^2 (*df*)	26.95 (2)***	1.03	30.08 (2)***
φ_c	0.22		0.27
Original CRS			
Minimum	24.91 (66/265)	30.56 (11/36)	24.02 (55/229)
Medium	40.80 (102/250)	35.05 (34/97)	44.44 (68/153)
Maximum	42.31 (11/26)	40.00 (4/10)	43.75 (7/16)
χ^2 (*df*)	15.73 (2)**	0.39 (2)	18.33 (2)***
φ^c	0.17		0.21

df = degrees of freedom
*p < 0.05   **p < 0.001   ***p < 0.0001

Similarly, the reweighted scale produces acceptable AUC values for the group, as a whole, and for non-Aboriginal women (Table 17). These values are increased from the AUCs produced by the original CRS. However, in contrast to AUC values associated with institutional incidents, the original CRS also produced acceptable AUCs values for these two groups. Also contrasting with the results for institutional adjustments, the reweighted scale is a significantly stronger predictor of institutional charges than is the original scale (χ^2(2, n= 541) = 3.87, p < 0.05).[15]

The pattern of AUC values produced for Aboriginal women contrast with the trend toward improved predictive validity found in the institutional incidents analyses (Table 14) as well as the other two analyses (Tables 16 and 18) pertaining to institutional charges. That is, the AUCs for the reweighted CRS were lower than the AUCs for the original CRS. On the one hand, although this drop was very minor (i.e., from 0.63 to 0.61), it is notable because it points to a slightly weakened prediction of institutional charges when using the reweighted CRS among Aboriginal women. On the other hand, both AUC values met acceptable levels and were higher than the equivalent analyses performed for predicting institutional incidents among Aboriginal women.

Table 17

Area Under the Receiver Operating Cure (AUC) for the Prediction of Institutional Charges using the CRS (Original and Reweighted)

	Overall	Aboriginal	Non-Aboriginal
Reweighted CRS: AUC (95% CI)	0.67 (0.60-0.75)	0.61 (0.48-0.74)	0.68 (0.58-0.78)
Original CRS: AUC (95% CI)	0.59 (0.51-0.68)	0.63 (0.49-0.78)	0.61 (0.50-0.71)

AUC = Area under the ROC Curve, CI = Confidence Interval

Also consistent with earlier results, survival analyses demonstrated that the original CRS and the reweighted CRS were significantly associated with time in both the overall sample and the non-Aboriginal sample (Table 18). Again, the level of significance increased with the reweighted scale. As with previous analyses, significant results were not found with the

---

[15] Results for non-Aboriginal (χ^2 (2, n = 399) = 2.22, p > 0.05) and Aboriginal (χ^2 (2, n = 143) = 0.04, p > 0.05) women were not significant.

Aboriginal sub-sample.

Although it is not central to the issue of predictive validity, it is interesting to compare the average uncensored survival days for institutional charges to institutional incidents. On average, charges occur much earlier than incidents. This may reflect patterns of recourse selected in the institutional environment. It may also provide insight as to why AUCs produced using institutional charges demonstrate greater predictive validity than institutional incidents, given the three month cut-off decision (i.e., a relatively large proportion of incidents may occur after the three month period).

Table 18

Survival Time to Charge, Overall and by Ethnicity, using the CRS (Original and Reweighted)

	N	% failed (n)	Test of Equality χ^2 (df)	Uncensored M (SE)	Censored M (SE)
Overall					
Reweighted					
Minimum	266	22.56 (60)		73.90 (9.66)	148.10 (8.08)
Medium	243	42.39 (103)	41.33 (2)***	73.48 (7.49)	104.80 (9.07)
Maximum	32	50.00 (16)		52.88 (19.28)	58.75 (20.79)
Original					
Minimum	265	24.71 (66)		57.79 (5.50)	139.70 (8.31)
Medium	250	40.80 (102)	15.78 (2)**	82.47 (8.70)	116.60 (8.94)
Maximum	26	42.31 (11)		56.55 (26.88)	70.20 (25.78)
Aboriginal					
Reweighted					
Minimum	39	28.21 (11)		58.64 (23.99)	156.30 (22.10)
Medium	92	35.87 (33)	3.71 (2)	78.03 (14.01)	103.90 (15.13)
Maximum	12	41.67 (5)		41.40 (22.45)	29.86 (23.22)
Original					
Minimum	36	30.56 (11)		37.18 (11.28)	139.20 (25.81)
Medium	97	35.05 (34)	3.08 (2)	84.76 (14.70)	112.10 (14.31)
Maximum	10	40.00 (4)		34.00 (24.84)	29.00 (26.87)
Non-Aboriginal					
Reweighted					
Minimum	227	21.59 (49)		77.33 (10.59)	146.80 (8.70)
Medium	151	46.36 (70)	42.04 (2)***	71.33 (8.89)	105.40 (11.24)
Maximum	20	55.00 (11)		58.09 (26.66)	81.22 (31.34)
Original					
Minimum	229	24.02 (55)		61.91 (6.08)	139.80 (8.78)
Medium	153	44.44 (68)	16.27 (2) **	81.32 (10.85)	120.00 (11.45)
Maximum	16	43.75 (7)		69.43 (40.51)	97.67 (37.45)

df = degrees of freedom, *M* = Mean, *SE* = Standard Error, *n* = sample size, uncensored *M* = the average number of day to have an event, censored *M* = average number of days at risk before period ended
*$p < 0.05$ **$p < 0.001$ ***$p < 0.0001$

Main Analyses: Validation Sample

With evidence that reweighting improved the predictive validity of the CRS in hand, the next portion of analyses were designed to validate whether applying the same weighting to a distinct subsample of participants will reproduce similar improvements.

Classification across Ethnicity

As demonstrated with the construction sample, the original CRS and final decision tend to place proportionally more Aboriginal women in higher security levels relative to the reweighted CRS (Table 19). In this case, the association between ethnicity and offender security level was strongest for the original CRS recommendation, rather than the final decision. Consistent with earlier results, the reweighted scale still produces a significant chi-square. Therefore, even when reweighting in undertaken to apply greater weighting to items that are relevant to the prediction of institutional adjustment, the majority of Aboriginal women would be recommended to medium security while the majority of non-Aboriginal women would be recommended to minimum security. Therefore, over-representation accounts for some of the discrepancy in security level recommendations and placements across ethnicity.

Table 19

Associations between Security Level Classifications and Ethnicity

		Aboriginal %	Non-Aboriginal %	χ^2 (*df*)	φ_c
Based on the Reweighted CRS	Minimum	30.77	59.40		
	Medium	60.84	35.84	34.57 (2) ***	0.25
	Maximum	8.39	4.76		
Based on the Original CRS	Minimum	27.97	63.16		
	Medium	61.54	33.83	55.66 (2) ***	0.32
	Maximum	10.49	3.02		
Final Decision	Minimum	22.38	54.89		
	Medium	72.73	40.85	45.83 (2) ***	0.29
	Maximum	4.90	4.26		

df = degrees of freedom
*$p < 0.05$ **$p < 0.001$ ***$p < 0.0001$

Concordance Rates

The concordance rates obtained for the validation sample (Table 20) were within a similar range as those found with the construction sample (Table 11). There are, however, some notable differences. First, the overall concordance between the final decision and the reweighted CRS (71.96%) is greater than the concordance between the final decision and the original CRS (70.85%). The reverse of this pattern was found earlier. Previously, it was proposed that the pattern of results found here may suggest that there is value added by the reweighting such that its recommendations give greater weight to those aspects that are important to determining security level placements. However, even if this proposition were true, these percentages only differ slightly and would only apply to non-Aboriginal women, in this case. The concordance rates between the final decision and both the original (65.04%) and reweighted CRS (62.94%) were low for Aboriginal women. In the case of the former, placement in higher or lower levels are equivalent (17.48%). For the latter, there is a greater tendency for the reweighted scale to classify women in lower security levels (20.98%) rather than classify in higher security levels (16.08%), relative to the final decision.

Table 20

Percentage of Concordance, Lower Classification and Higher Classification

		Concordance %	Lower Classification %	Higher Classification %
Between Original CRS and Reweighted CRS	Overall Percentages	73.98	11.63	14.40
	Aboriginal Percentages	76.92	13.99	9.09
	Non-Aboriginal Percentages	72.93	10.78	16.30
Between Final Decision and Reweighted CRS	Overall Percentages	71.96	16.04	11.99
	Aboriginal Percentages	62.94	20.98	16.08
	Non-Aboriginal Percentages	75.20	14.28	10.53
Between Original CRS and Final Decision	Overall Percentages	70.85	17.90	11.25
	Aboriginal Percentages	65.04	17.48	17.48
	Non-Aboriginal Percentages	72.93	18.05	9.02

Convergent Validity

In all cases, except one, the association between the reweighted OSL recommendations and measures of dynamic and static risk were significant (Table 21). In line with findings of the construction sample, these results demonstrate that risk increased as recommended security level increased. The relationship between recommended CRS level and dynamic risk was not significant for Aboriginal women ($p = 0.40$). This non-significant finding may, in part, be due to the high number of Aboriginal women with a high level of dynamic risk who would be recommended for minimum security (56.76%). Comparatively, only 5.41% of Aboriginal women recommended for minimum security had a low level of dynamic risk. As well, a greater percentage of Aboriginal women at medium security were rated as having high dynamic risk than Aboriginal women at maximum security. The incongruence of these results is likely related to three issues: (1) very few Aboriginal women have low dynamic risk ratings (3 out of 129 women), (2) when risk does not translate into poor institutional adjustment, a reweighted scale may recommend placement of women with high risk levels in low security levels and (3) smaller sample sizes of Aboriginal women require the use of Fisher's exact test, which tends to be more

conservative (i.e., less likely to find significant results). Small sample sizes also produce percentages that can vary widely, despite only small changes in raw scores.

Table 21

Associations between Security Levels (Recommendations and Placement) and Measures of Risk and Reintegration Potential

		Dynamic Risk			Static Risk		
		Low % (n)	Medium % (n)	High % (n)	Low % (n)	Medium % (n)	High % (n)
Reweighted CRS: Overall	Minimum	20.55 (52)	54.15 (137)	25.30 (64)	49.64 (139)	37.50 (105)	12.86 (36)
	Medium	1.46 (3)	35.44 (73)	63.11 (130)	20.09 (46)	55.90 (128)	24.02 (55)
	Maximum	0 (0)	23.33 (7)	76.67 (23)	6.45 (2)	35.48 (11)	58.06 (18)
χ^2 (df)		96.65 (4) ***			80.59 (4) ***		
φ_c		0.31			0.27		
Reweighted CRS: Aboriginal	Minimum	5.41 (2)	37.84 (14)	56.76 (21)	36.36 (16)	43.18 (19)	20.45 (9)
	Medium	1.25 (1)	27.50 (22)	71.25 (57)	12.64 (11)	58.62 (51)	28.74 (25)
	Maximum	0 (0)	33.33 (4)	66.67 (8)	8.33 (1)	25.00 (3)	66.67 (8)
Fisher's Exact		p = 0.40			p = 0.001**		
Reweighted CRS: Non-Aboriginal	Minimum	23.15 (50)	56.94 (123)	19.91 (43)	52.12 (123)	36.44 (86)	11.44 (27)
	Medium	1.59 (2)	40.48 (51)	57.94 (73)	24.65 (35)	54.23 (77)	21.13 (30)
	Maximum	0 (0)	16.67 (3)	83.33 (15)	5.26 (1)	42.11 (8)	52.63 (10)
χ^2 (df)		79.91 (4)***			49.29 (4)***		
φ_c		0.33			0.25		

df = degrees of freedom, n = sample size

*p < 0.05   **p < 0.001   ***p < 0.0001

Predictive Validity

The same three methods, used with the construction sample, were used to test the predictive validity of the reweighted CRS with validation sample data. Below, chi-square tests of independence, AUCs and survival analyses are presented to examine the CRS's (original and reweighted) ability to predict outcome variables, institutional incidents and charges.

Institutional Incidents

As with the construction sample, there was a trend of increasing association between institutional incident involvement and recommended CRS when using the reweighted CRS, relative to the original CRS (Table 22). One clear difference from the construction sample is finding significant associations for Aboriginal women. In the construction sample, there were no associations between incident involvement and recommended security level for both the original and reweighted CRS. In the present validation sample, the strength of the association between these variables was greatest for Aboriginal women. The fluctuation in these results appears to be due to relatively small differences in the raw number of women involved in institutional incidents. For example, whereas 5/12 (41.67%) Aboriginal women recommended for maximum security (by the reweighted CRS) in the construction had an institutional incident, 10 out of 12 (83.33%) Aboriginal women recommended for maximum security had an institutional incident in the validation sample. Hence, these differences reflect a small sample size, whereby each individual woman has a large impact on the likelihood of finding significance.

Table 22

Association between the CRS (Original and Reweighted) and Institutional Incidents

	Cases Involving Institutional Incidents %		
	Overall	**Aboriginal**	**Non-Aboriginal**
Reweighted CRS			
Minimum	21.00 (59/281)	20.45 (9/44)	21.10 (50/237)
Medium	35.65 (82/230)	24.14 (21/87)	42.66 (61/143)
Maximum	58.06 (18/31)	83.33 (10/12)	42.11 (8/19)
χ^2 (*df*)	26.20 (2)***	20.12 (2)***	21.25 (2)***
φ_c	0.22	0.38	0.23
Original CRS			
Minimum	22.60 (66/292)	17.50 (7/40)	23.41 (59/252)
Medium	34.98 (78/223)	26.14 (23/88)	40.74 (55/135)
Maximum	55.56 (15/27)	66.67 (10/15)	41.67 (5/12)
χ^2 (*df*)	18.76 (2)***	13.47 (2) *	13.44 (2) *
φ_c	0.19	0.31	0.18

df = degrees of freedom
*$p < 0.05$ **$p < 0.001$ ***$p < 0.0001$

    The proportion of women in the validation sample observed for three months following the date of the CRS decision was very similar to the proportions found among the construction sample. Overall, 60.89% (330/544) of offenders remained in the institution for three months following their OSL decision; by ethnicity, these numbers were 56.64% (81/143) and 62.41% (249/399) of Aboriginal and non-Aboriginal women, respectively.

    In general, the AUCs demonstrated stronger prediction of incident occurrence in the validation sample (Table 23) than the construction sample (Table 14). This is particularly true for Aboriginal women, who went from having the lowest AUCs in the construction sample (reweighted CRS AUC = 0.58) to the highest AUCs in the validation sample (reweighted CRS AUC = 0.66). As discussed earlier, the volatility of this information can be attributed to the small sample of Aboriginal women. Moreover, the small sample sizes often result in large confidence intervals (e.g., 0.50-0.81),

which reflect the large amount of error associated with the AUC value and indicate that caution should be taken when interpreting these findings.

Aside from the AUCs obtained for Aboriginal women, there is a slight improvement in the prediction of incident involvement when using the reweighted CRS, as compared to the original CRS. However, non-significant chi-square for the ROC contrast estimation (χ^2 (2, $n = 542$) = 0.41, p > 0.05) indicate that there is not a significant difference between these two measures.[16]

Table 23

AUCs for the Prediction of Institutional Incidents using the CRS (Original and Reweighted)

	Overall	Aboriginal	Non-Aboriginal
Reweighted CRS AUC (95% CI)	0.62 (0.55-0.69)	0.66 (0.50-0.81)	0.62 (0.54-0.70)
Original CRS AUC (95% CI)	0.60 (0.53-0.67)	0.67 (0.53-0.82)	0.59 (0.51-0.67)

AUC = Area under the ROC Curve, CI = Confidence Interval

Across all three survival analyses (Table 24), the original CRS and the reweighted CRS were significantly associated with time to incident. Consistent with the findings for the construction sample, associations are stronger when using recommendations based on the reweighted scale rather than the original scale.

There are some discrepancies between the validation sample and construction sample. The most salient distinction is the finding of significant results among Aboriginal women in the current sample. This shift is consistent with the two previous sets of analyses, all of which occur as a result of small changes in incident involvement that become magnified by the fact that the sub-sample of Aboriginal women is small.

---

[16] Results for non-Aboriginal (χ^2 (2) = 0.63, $p > 0.05$) and Aboriginal (χ^2 (2) = 0.09, $p > 0.05$) women were also not significant.

Table 24

Survival Time to Incident, Overall and by Ethnicity, using the CRS (Original and Reweighted)

	N	% failed (*n*)	Test of Equality χ^2 (*df*)	Uncensored *M* (*SE*)	Censored *M* (*SE*)
Overall					
Reweighted					
Minimum	281	21.00 (59)		92.58 (10.98)	148.20 (8.62)
Medium	230	35.65 (82)	34.54 (2)***	83.60 (9.30)	106.10 (8.39)
Maximum	31	58.06 (18)		63.50 (15.41)	122.30 (44.03)
Original					
Minimum	292	22.60 (66)		92.14 (11.03)	135.90 (8.03)
Medium	223	34.98 (78)	23.28 (2)***	81.15 (8.88)	127.40 (10.00)
Maximum	27	55.56 (15)		69.93 (17.87)	84.42 (42.38)
Aboriginal					
Reweighted					
Minimum	44	20.45 (9)		59.00 (21.49)	166.70 (24.34)
Medium	87	24.14 (21)	20.81 (2)***	79.95 (13.36)	104.70 (12.58)
Maximum	12	83.33 (10)		69.30 (24.16)	70.71 (50.00)
Original					
Minimum	40	17.50 (7)		83.13 (29.62)	142.00 (23.25)
Medium	88	26.14 (23)	18.15 (2)***	70.78 (12.04)	123.20 (14.36)
Maximum	15	66.66 (10)		69.30 (24.16)	34.80 (16.15)
Non-Aboriginal					
Reweighted					
Minimum	237	21.10 (50)		98.62 (12.24)	144.80 (9.17)
Medium	143	42.66 (61)	27.90 (2)***	84.84 (11.68)	107.30 (11.33)
Maximum	19	42.10 (8)		56.25 (18.62)	133.30 (51.24)
Original					
Minimum	252	23.41 (59)		93.20 (11.90)	134.90 (8.54)
Medium	135	40.74 (55)	13.27 (2)***	85.49 (11.56)	130.80 (13.94)
Maximum	12	41.67 (5)		71.20 (26.94)	119.90 (70.84)

df = degrees of freedom, *M* = Mean, *SE* = Standard Error, *n* = sample size, uncensored *M* = the average number of day to have an event, censored *M* = average number of days at risk before period ended

*$p < 0.05$ **$p < 0.001$ ***$p < 0.0001$

Institutional Charges

The validation sample results of the chi-square analyses for the prediction of institutional charges mirror those of the construction sample. The overall sample and non-Aboriginal sub-sample are significantly associated with OSL recommendations given by the original and reweighted CRS (Table 25). However, these associations are stronger with the reweighted scale.

Table 25

Association between the CRS (Original and Reweighted) and Institutional Charges

	Cases Involved in Institutional Charges %		
	Overall	**Aboriginal**	**Non-Aboriginal**
Reweighted CRS			
Minimum	25.98 (73/281)	34.09 (15/44)	24.47 (58/237)
Medium	46.52 (107/230)	47.13 (40/87)	46.15 (66/143)
Maximum	51.61 (16/31)	50.00 (6/12)	52.63 (10/19)
χ^2 (*df*)	26.52 (2)***	2.26 (2)	22.04 (2)***
φ_c	0.22		0.24
Original CRS			
Minimum	27.40 (80/292)	40.00 (16/40)	25.50 (64/252)
Medium	46.19 (103/223)	44.32 (39/88)	47.41 (64/135)
Maximum	48.15 (13/27)	46.67 (7/15)	50.00 (6/12)
χ^2 (*df*)	21.11 (2)***	0.28 (2)	20.59 (2)***
φ_c	0.20		0.23

df = degrees of freedom
*$p < 0.05$ **$p < 0.001$ ***$p < 0.0001$

The AUC values obtained for the prediction of institutional charges (Table 26) are virtually identical to the values obtained with the construction sample (Table 17). Consistent with the results of the construction sample, the reweighted scale is a stronger predictor of institutional charges than is the original scale (χ^2 (2, $n = 542$) = 3.58, p < 0.06).[17] However, this difference is only marginally

---

[17] Results for non-Aboriginal (χ^2 (2, $n = 398$) = 2.82, $p > 0.05$) and Aboriginal (χ^2 (2, $n = 143$) = 0.39, $p > 0.05$) women were not significant.

significant.

The results for Aboriginal women are distinctive from the construction sample results in that they are lower and below the 0.60 cut-off to delineate acceptable predictive capacity. Consistent with earlier discussion of fluctuations found among the results for Aboriginal women, these variations between sub-samples and large confidence intervals reflect a small sample size and should be interpreted cautiously.

Table 26

Area Under the Receiver Operating Cure (AUC) for the Prediction of Institutional Charges using the CRS (Original and Reweighted)

	Overall	**Aboriginal**	**Non-Aboriginal**
Reweighted CRS			
AUC (95% CI)	0.66 (0.60-0.71)	0.59 (0.47-0.70)	0.67 (0.60-0.74)
Original CRS			
AUC (95% CI)	0.60 (0.54-0.65)	0.53 (0.42-0.65)	0.61 (0.54-0.67)

AUC = Area under the ROC Curve, CI = Confidence Interval

The pattern of results for the time to charge (Table 27) are similar to those found for incidents, above (Table 24), and to survival time to charge for the construction sample (Table 18). The main difference is that, in the current analyses, there is not a significant relationship between recommended security level and time to charge for Aboriginal women. In general, the survival functions for Aboriginal women have been highly variable between sub-samples (construction sample versus validation sample) and between institutional adjustment outcome (institutional incidents versus institutional charges). This variability is likely influenced by the sample size in so far as small changes in incident or charge involvement can have large influences on calculations used to assess significance.

Table 27

Survival Time to Charge for Overall Sample using the CRS (Original and Reweighted)

	N	% failed (*n*)	Test of Equality χ^2 (*df*)	Uncensored *M* (*SE*)	Censored *M* (*SE*)
Overall					
Reweighted					
Minimum	281	25.98 (73)		90.60 (10.19)	128.50 (6.69)
Medium	230	46.52 (107)	52.25 (2)***	52.85 (4.89)	87.98 (8.37)
Maximum	31	51.61 (16)		47.25 (13.89)	130.40 (41.28)
Original					
Minimum	292	27.40 (80)		73.39 (8.91)	119.70 (6.11)
Medium	223	46.19 (103)	27.48 (2)***	64.17 (6.06)	103.90 (9.93)
Maximum	27	48.15 (13)		41.85 (14.48)	119.00 (45.87)
Aboriginal					
Reweighted					
Minimum	44	34.09 (15)		62.40 (20.39)	139.90 (20.02)
Medium	87	47.13 (41)	4.72 (2)***	61.95 (8.16)	78.20 (13.22)
Maximum	12	50.00 (6)		13.33 (3.47)	166.30 (63.43)
Original					
Minimum	40	40.00 (16)		49.75 (17.90)	117.50 (17.19)
Medium	88	44.32 (39)	3.76 (2)	68.15 (8.78)	99.29 (15.52)
Maximum	15	46.66 (7)		14.57 (3.18)	120.80 (53.77)
Non-Aboriginal					
Reweighted					
Minimum	237	24.47 (58)		97.90 (11.59)	126.70 (7.07)
Medium	143	46.15 (66)	48.22 (2)***	47.20 (6.04)	93.82 (10.79)
Maximum	19	52.63 (10)		67.60 (19.70)	106.40 (55.80)
Original					
Minimum	252	25.40 (64)		79.30 (10.12)	120.00 (6.54)
Medium	135	47.41 (64)	26.36 (2)***	61.75 (8.19)	107.10 (12.99)
Maximum	12	50.00 (6)		73.67 (26.46)	116.70 (83.23)

df = degrees of freedom, *M* = Mean, *SE* = Standard Error, *n* = sample size, uncensored *M* = the average number of day to have an event, censored *M* = average number of days at risk before period ended
*$p < 0.05$ **$p < 0.001$ ***$p < 0.0001$

From a broader perspective, smaller sample sizes among Aboriginal women resulted in their lesser influence on the initial reweighting exercise. Therefore, relationships between the reweighted scale and outcome measures are more tenuous for Aboriginal women.

<center>**Discussion**</center>

The goal of reweighting the Custody Rating Scale was to improve the scale in two ways: (1) to address the disproportionate number of Aboriginal women at higher security levels and (2) to improve the validity of the scale for women offenders. The results of the analyses pertaining to the construction and validation samples demonstrate that these two goals are partially achieved through the reweighting.

Proportion of Aboriginal Women at Higher Security Levels

When examining classifications across ethnicity, it appears that the reweighting improves the disproportionality in the distribution of Aboriginal women at higher security levels. Reweighting enables those aspects that are more relevant to the prediction of institutional outcomes to have greater influence over the recommended security level classification than would an unweighted scale. This reduces the propensity to place Aboriginal women in higher security levels if the risks they pose do not result in institutional misconducts (i.e., incidents or charges). Despite this improvement, there remains a significantly greater number of Aboriginal women at higher levels of security than non-Aboriginal women. This supports the proposition that much of the reason for disproportional numbers of Aboriginal women at higher security levels is due to their over-representation in higher risk categories. Future research, with a scale that demonstrates consistent predictive validity for Aboriginal women, could re-examine this issue and would have the capacity to draw more definitive conclusions on the issue of over-representation versus over-classification.

With the current preliminary evidence that over-representation will continue to occur after reweighting is applied, the question remains: to what extent should an initial custody classification measure be designed to underplay risks within the institution and to the public in order to ensure proportional representation across security levels? Perhaps the emphasis should be on the reduction of these risks in subsequent security level assessments and over the course of the sentence rather than proportional initial custody classification. Future research could explore which case management approaches are most effective in reducing the security levels placements of women across ethnicity.

Validity of the Reweighted CRS for Women

Reweighting the CRS produced a consistent pattern of improved predictive validity for women overall and for non-Aboriginal women. This means that for these two (non-distinct) groups, associations of the CRS to institutional incidents and outcomes were increased by reweighting.

<center>54</center>

Although these improvements did not always represent a significant change, on some occasions reweighting altered results enough to produce "acceptable" values when they were originally lower. For example AUC values for the prediction of institutional charges for the overall construction sample went from below acceptable, when using the original CRS, to between acceptable and good, when using the reweighted CRS.

However, results were not consistent across all groups. Hannah-Moffat and Shaw (2001) argue that the population of women in federal prison is much smaller than the population of men, less involved in criminal activity and more heterogeneous. In this case, the heterogeneity and size of the population limited the extent of improvements that were produced by a reweighting. Namely, the small number of women in the Aboriginal sub-samples had an influence on two issues, evident throughout the construction and validation sample. First, the relative size of the Aboriginal sample to the non-Aboriginal sample meant that the reweighting of the CRS reflected the associations between items and institutional charges for non-Aboriginal women. In other words, in a situation where an item was significant for one group, but not the other, the likelihood is that new weights were representative of the pattern present for non-Aboriginal women. Given the heterogeneity of Aboriginal and non-Aboriginal sub-samples, it is likely that weights representative of patterns present for non-Aboriginal women produced a reweighted measure that is less valid for Aboriginal women. The small sample size of Aboriginal women can also create large fluctuations in overall percentages when, in actuality, raw numbers have not changed substantially. For example, it would seem that Aboriginal women in maximum security had far more incidents involvement when comparing the percentages across the validation (83.33%) and construction (41.66%) samples. However, this difference is only due to an additional five participants being involved in incidents in the validation sample.

Alternative Approaches

One way of dealing with the low influence of Aboriginal women on the reweighting of the CRS would be to produce separately weighted scales, by ethnicity. This option was not pursued here for two reasons. First, a decision was made at the onset of the project to produce one reweighted scale to simplify the implementation of a reweighted scale into daily operations. Second, the small sample size of Aboriginal women could produce highly variable weightings that are not generalizable to the population.

An alternative to separate reweightings would be to produce a gender-informed scale specifically designed for women. Although increases in the predictive validity for the overall and non-Aboriginal samples demonstrated positive progression in the validity of the CRS, questions remain about the substantive improvement of the scale following a reweighting. For example, the low probability of involvement in an institutional charge among those who would be recommended to maximum security (Appendix D) indicates that many women may be inappropriately placed. These low probability values are also reflected in the findings of "acceptable" predictive validity. Although prediction of institutional adjustment problems becomes better than chance alone, "acceptable" results signify that the reweighting is performing below "good" and is limited in predictive validity. As a result, many women can be classified to higher or lower security level relative to their subsequent institutional adjustment behaviours.

In the context of the current study, it is likely that the limited predictive validity of the reweighted CRS for women can be traced back to preliminary results of the reweighting, which demonstrate that very few items in the reweighted CRS are significant predictors of institutional adjustment outcomes. This finding echoes those of Blanchette et al. (2002) and Barnum & Gobeil (in press), where items of the original CRS demonstrated weak associations with institutional adjustment outcomes. In contrast, associations between these items and institutional adjustment outcomes are more consistent for men (Gobeil, 2011). This would suggest that the scale and/or the outcomes being examined have lower relevance for women than they do for men. It suggests that we need to re-evaluate our outcome measures to ensure they are the most relevant when developing a scale.

Future research on the feasibility of a gender-informed instrument could explore which items and outcomes would best predict incarceration behaviours for women. For example, Farr (2000) suggests that because women pose little institutional risk, other outcomes (e.g., treatment needs) would be more suitable for classification than those that are based on risk. Monture-Angus (2000) argues that the notion of "risk" is not compatible with Aboriginal culture, law and traditions. Further, she suggests that measures designed to assess risk are problematic for Aboriginal offenders who, as a result of negative impacts from colonialism, may be rated more negatively on certain portions of the assessment. Hence, the CRS may include elements that demonstrate differential item functioning[18] (Zumbo, 1999) that would automatically create scoring biases.

---

[18] "Differential Item Functioning occurs when examinees from different groups show different probabilities of success on (or endorsing) the item *after matching on the underlying ability [or, in this case, underlying likelihood of institutional misconduct]* that the item is intended to measure" (Zumbo, 1999; p. 12).

Research has also demonstrated unique factors that lead toward recidivism for women, which are not typically observed with men. A gender-informed instrument could explore how factors such as unhealthy relationships, trauma, mental illness, substance abuse (Hardyman & Van Coorhis, 2004; Salisbury & Van Voorhis, 2009) and family contact (Blanchette, 2005 as cited in Blanchette & Brown, 2006) affect institutional adjustment. Further, this type of re-imagining of an initial custody classification instrument for women could allow for greater sensitivity to ethnic differences that are important to institutional adjustment for both Aboriginal and non-Aboriginal women.

Recently, there have been significant advancements in the construction of gender-informed tools within CSC. The Custody Rating Scale (CRS) is used in determining initial security placements, however subsequent security classification reviews are completed with the Security Reclassification Scale for Women (Blanchette & Taylor, 2005). The SRSW determines a security level for offenders (i.e., minimum, medium, maximum) through the scaling of nine dynamic factors[19] thereby allowing for re-assessment during the offenders' incarceration. Since its national implementation in 2005, the scale was revalidated in 2008 (Gobeil, 2008) to ensure its continued reliability and validity for review periods of at least six months.

Examining *risk* more generally, CSC Research Branch attempted to develop, from the ground-up, a gender-informed, dynamic risk assessment scale that could be used to predict risk of recidivism in federal women offenders. Unfortunately, a gender-informed *dynamic* risk scale for women could not be validated. Variables that contributed to the model were largely *static* in nature (i.e., not amenable to change) and none of the gender-informed factors (e.g., self-esteem, victimization) contributed to increasing the scale's predictive accuracy. In this case, the gender-informed variables did not increase ability to predict risk of recidivism (Zakaria, Allenby, Derkzen, & Jones, 2013).

Conclusion

Overall, reweighting the CRS items slightly improved the scale for non-Aboriginal women, but had less of an effect for Aboriginal women. Given these differing patterns and the limited magnitude of improvements, the implementation of the reweighted scale is not recommended at this

---

[19] The SRSW determines a security level for offenders (i.e., minimum, medium, maximum) through the scaling of nine dynamic factors: 1) placement in involuntary segregation, 2) progress or motivation regarding correctional plan, 3) presence of serious disciplinary offences, 4) number of recorded incidents, 5) number of successful escorted temporary absences, 6) CRS history rating scale, 7) most recent level of pay, 8) ever unlawfully at large (UAL) from temporary absence, work release, or supervision, and 9) family contact. Each variable is optimally weighted through statistical procedures.

time. Questions remain regarding the best assessment approaches for women.   These questions are particularly relevant for Aboriginal women, whose relatively small numbers and over-representation at higher security levels create challenges in assuring that a valid instrument is available for their classification. Ultimately, finding solutions to this challenge is central to respecting the spirit of the CCRA and will minimize risks to the public, staff members and women offenders.

References

Barnum, G. & Gobeil, R. (2012). *Revalidation of the custody rating scale for Aboriginal and non-Aboriginal women*. Ottawa, ON: Correctional Service of Canada.

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, England: John Wiley & Sons.

Blanchette, K. & Motiuk, L.L. (2004). Taking down the straw man: A reply to Webster and Doob. *Canadian Journal of Criminology and Criminal Justice*, *46*, 621-630.

Blanchette, K. & Taylor, K. (2005). *Development and field-test of a gender-informed security reclassification scale for women offenders*. (Report No. R-167). Ottawa, ON: Correctional Service of Canada.

Blanchette, K., Verbrugge, P. & Wichmann, C. (2002). *The custody rating scale, initial security placement, and women offenders* (Report No. R-127). Ottawa, ON: Correctional Service of Canada. Retrieved from http://www.csc-scc.gc.ca/text/rsrch/reports/r127/r127_e.pdf.

Bonta, J. & Motiuk, L.L. (1992). Inmate classification. *Journal of Criminal Justice*, *20*, 343-353.

Brennan, T., & Austin, J. (1997). *Women in jail: Classification issues* (NIC Accession No. 013768). Washington, DC: National Institute of Corrections. Retrieved from http://nicic.gov/pubs/1997/013768.pdf.

Canadian Human Rights Commission. (2003). *Protecting their rights: A systemic review of human rights in correctional services for federally sentenced women*. Ottawa, ON: Canadian Human Rights Commission.

Correctional Service of Canada (2010). *Commissioner's Directive 705-7*: Security classification and penitentiary placement. Ottawa, ON: Correctional Service Canada. http://www.csc-scc.gc.ca/text/plcy/cdshtm/705-7-cd-eng.shtml

Corrections and Conditional Release Act [CCRA]. (1992). Retrieved from http://laws-lois.justice.gc.ca/eng/acts/C-44.6/

Coxe, S., West, S.G. & Aiken, L.S. (2009). The analysis of count data: A gentle introduction to Poisson Regression and its alternatives. *Journal of Personality Assessment, 91,* 121-136.

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

Gallop, R. J. (2010). Determination and interpretation of the OOP for ROC's with PROC LOGISTIC. Retrieved from: http://www.nesug.org/proceedings/nesug01/st/st9006r

Gobeil, R. (2008). Assessing security reclassification with male Aboriginal and non-Aboriginal offenders (Report No. R-203). Ottawa, ON: Correctional Service of Canada.

Gobeil, R. (2011). The Custody Rating Scale as applied to male offenders. (Report No. R-257). Ottawa, ON: Correctional Service of Canada.

Grant, B.A. & Luciani, F. (1998). *Security classification using the custody rating scale* (Report No. R-67). Ottawa, ON: Correctional Service of Canada.

Hannah-Moffat, K. & Shaw, M. (2001). Taking risks: Incorporating gender and culture into the classification and assessment of federally sentenced women in Canada. Ottawa, Ontario: Status of Women Canada.

Hardyman, P.L. (2001). *Validation and refinement of objective prison classification systems for women: The experience of four states and common themes*. Washington, DC: The Institute on Crime, Justice and Corrections, National Institute of Corrections.

Hardyman, P.L. & Van Voorhis, P. (2004). *Developing gender-specific classification systems for women offenders*. National Institute of Correction (NIC Accession Number 018931).

Lin, M.J. (2006). Modifying Danger Assessment (DA) Scale for better performance on predicting lethal risk among male intimate users through weighting procedures. *Asian Journal of Domestic Violence and Sexual Offense*, *2*, 45-63.

Luciani, F. P., Motiuk, L. L., & Nafekh, M. (1996). *An operational review of the custody rating scale: Reliability, validity and practical utility* (Report R-47). Research Division, Correctional Service of Canada, Ottawa, Canada.

Monture-Angus, P. (2000). Aboriginal women and correctional practice: Reflections on the Task Force on Federally Sentence Women. In *An Ideal Prison? Critical Essays on Women's*

Imprisonment in Canada. K. Hannah-Moffat and M. Shaw, (Eds.). Halifax: Fernwood Publishing.

Nuffield, J. (1982). Parole decision-making in Canada: Research toward decision guidelines. Ottawa, ON: Communications Division; Supply and Services Canada.

Porporino, F.J., Juciani, F., Motiuk, L., Johnston, M. & Mainwaring, B. (1989). *Pilot implementation of a custody rating scale: Interim report* (Report No. R-02). Ottawa, ON: Correctional Service of Canada.

Raftery, A.E. (1995). Bayesian model selection in social research. *Sociological methodology*, *25*, 111-163.

Salisbury, E.J. & Van Voorhis, P. (2009). Gendered Pathways: A quantitative investigation of women probationers' path to incarceration. *Criminal Justice and Behavior, 36*, 541-566.

Swets, J.A. (1988). Measuring the accuracy of diagnostic systems. *Science, 240*, 1285-1293.

Verbrugge & Blanchette (2002). *The Validity of the Custody Rating Scale for the Initial Security Classification of Aboriginal Women*, in Forum on Corrections Research, Volume 14, Number 3. Ottawa: Correctional Service of Canada.

Webster, C.M. & Doob, A.N. (2004). "Taking down the straw man" or building a house of straw? Validity, equity, and the Custody Rating Scale. *Canadian Journal of Criminology and Criminal Justice*, *46*, 395-422.

Zakaria, D., Allenby, K., Derkzen, D.M. & Jones, N. (2013) *Preliminary Development of a Dynamic Risk Assessment Tool for Women Offenders: An Examination of Gender-Neutral and Gender-Specific Variables.* (Report No. R-280). Ottawa, ON: Correctional Service of Canada.

Zumbo, B.D. (1999). *A handbook on theory and methods of differential item functioning (DIF)*. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.

Zumbo, B.D., Gadermann, A.M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for Likert rating scales. *Journal of Modern Applied Statistical Methods*, *6*, 21-29.

Appendices

Appendix A: Custody Rating Scale

Institutional Adjustment Subscale

Item	Points
1. History of involvement in institutional incidents	
a. no prior involvement	0
b. any prior involvement	2
c. prior involvement in one or more incidents in "greatest" or "high" severity categories	2
d. prior involvement during last five years of incarceration:	
▪ in an assault (no weapon or serious injury)	1
▪ in a riot or major disturbance	2
▪ in an assault (using a weapon or causing serious injury)	2
e. involvement in one or more serious incidents prior to sentencing and / or pending placement from current commitment	5
8 X Total of a. to e.	
2. Escape history	
a. no escapes or attempts	0
b. an escape or attempt from minimum or community custody with no actual or threatened violence:	
▪ over two years ago	4
▪ in last two years	12
c. an escape of attempt from medium or maximum custody or an escape from minimum or community custody with actual or threatened violence:	

▪ over two years ago	20
▪ in last two years	28
3. Street stability	
a. above average	0
b. average	16
c. below average	32
4. Alcohol / drug use	
a. no identifiable problems	0
b. abuse affecting one or more life areas	3
c. serious abuse affecting several life areas	6
5. Age (at the time of sentencing)	
a. 18 years or less	24
b. 19	22
c. 20	20
d. 21	18
e. 22	16
f. 23	14
g. 24	12
h. 25	10
i. 26	8
j. 27	6
k. 28	4
l. 29	2
m. 30 years or more	0

Security Risk Subscale

Item	Points
1. Number of prior convictions	
a. none	0
b. one	3
c. 2 to 4	6
d. 5 to 9	9
e. 10 to 14	12
f. over 15	15
2. Most serious outstanding charge	
a. no outstanding charges	0
b. minor	12
c. moderate	15
d. serious	25
e. major	35
3. Severity of current offence	
a. minor or moderate	12
b. serious or major	36
c. extreme	69
4. Sentence length	
a. 1 day to 4 years	5
b. 5 to 9 years	20
c. 10 to 24 years	45

d. over 24 years	65	
5. Street stability		
a. above average	0	
b. average	5	
c. below average	10	
6. Prior parole and / or statutory release		
a. none	0	
b. prior parole release	1 point for each	
c. prior statutory release	2 points for each	
7. Age		
a. 25 years or less	30	
b. 26	27	
c. 27	24	
d. 28	21	
e. 29	18	
f. 30	15	
g. 31	12	
h. 32	9	
i. 33	6	
j. 34	3	
k. 35 years or more	0	

Appendix B: Development of the Burgess Scale Based on Charges

CRS Items	Predictor Distributions (%)	Charge Rate per Predictor Category (%)	Intervals of a Full 5% Above or Below the Base Charge Rate (33.1%)	Final Score for the Category
1a. Involvement in prior incidents				
no	52.0	26.2	(26.2-33.1)/5 = -1.4	-1.4
yes	48.0	46.2	(46.2-33.1)/5 = 2.6	2.6
1b. Prior involvement in incidents in "greatest" or "high" severity category				
no	91.6	32.5	(32.5-33.1)/5 = -0.2	-0.2
yes	8.4	40.5	(40.5-33.1)/5 = 1.4	1.4
1c. Prior involvement during last five years of incarceration: - In an assault (no weapon or serious injury) - In a riot or major disturbance - In an assault (using a weapon or causing serious injury)				
no prior involvement during last 5 years of incarceration	78.8	30.8	(30.8-33.1)/5 = -0.4	-0.4
involvement in an assault (no weapon or serious injury)	18.4	43.4	(43.4-33.1)/5 = 2.0	2.0
either involvement in a riot or major disturbance or in an assault (using a weapon or causing serious injury)	2.8	71.4	(71.43-33.1)/5 = 7.6	7.6
1d. Involvement in one or more serious incidents prior to sentencing and / or pending placement from current commitment				
no	92.2	32.1	(32.1-33.1)/5 = -0.2	-0.2
yes	7.8	51.9	(51.9-33.1)/5 = 3.8	3.8
2. Escape history				
no escape or attempt	91.1	32.4	(32.4-33.1)/5 = -0.2	-0.2
an escape or attempt from minimum or community custody with no actual or threatened violence over two years ago	3.9	35.0	(35.0-33.1)/5 = 0.4	0.4
an escape or attempt from	3.4	50.0	(50.0-33.1)/5 = 3.4	3.4

minimum or community custody with no actual or threatened violence in last two years				
an escape or attempt from medium or maximum custody or an escape from minimum or community custody with actual or threatened violence over two years ago	0.0	0.0	(0-33.1)/5 = -6.6	-6.6
an escape or attempt from medium or maximum custody or an escape from minimum or community custody with actual or threatened violence in last two years OR two or more escapes from any level within the last five years	1.7	60.0	(60.0-33.1)/5 = 5.4	5.4
3. Street stability				
Above average	5.0	14.8	(14.8-33.1)/5 = -3.6	-3.6
Average	44.7	31.5	(31.5-33.1)/5 =-0.4	-0.4
Below average	50.3	39.8	(39.8-33.1)/5 = 1.4	1.4
4. Alcohol / drug use				
no identifiable problems	12.3	16.7	(16.7-33.1)/5 = -3.2	-3.2
abuse affecting one or more life areas	18.4	36.7	(36.7-33.1)/5 = 0.8	0.8
serious abuse affecting several life areas	69.3	38.9	(38.9-33.1)/5 = 1.2	1.2
5. Age (at time of sentencing) - **Initial**				
18 years or less	0.6	33.3	(33.3-33.1)/5 = 0.0	0.0
19	1.7	42.9	(42.9-33.1)/5 = 2.0	2.0
20	2.8	31.3	(31.3-33.1)/5 = -0.4	-0.4
21	3.4	50.0	(50.0-33.1)/5 = 3.4	3.4
22	1.1	14.3	(14.3-33.1)/5 = -3.8	-3.8
23	6.1	36.7	(36.7-33.1)/5 = 0.8	0.8
24	2.2	23.5	(23.5-33.1)/5 = -2.0	-2.0
25	6.1	52.4	(52.4-33.1)/5 = 3.8	3.8

26	7.8	50.0	(50.0-33.1)/5 = 3.4	3.4
27	5.0	37.5	(37.5-33.1)/5 = 0.8	0.8
28	5.6	43.5	(43.5-33.1)/5 = 2.0	2.0
29	2.8	23.8	(23.8-33.1)/5 = -1.8	-1.8
30 years or more	54.7	30.2	(30.2-33.1)/5 = -0.6	-0.6
5. Age (at time of sentencing) - Revised				
30 years of age or more	54.7	30.2	(30.2-33.1)/5 = -0.6	-0.6
29 years of age or less	45.3	37.5	(37.5-33.1)/5 = 0.8	0.8
6. Number of prior convictions				
None	21.8	26.7	(26.7-33.1)/5 = -1.2	-1.2
One	9.5	28.8	(28.8-33.1)/5 = -0.8	-0.8
Two to four	19.6	34.0	(34.0-33.1)/5 = 0.2	0.2
Five to nine	17.3	38.8	(38.8-33.1)/5 = 1.2	1.2
Ten to fourteen	8.9	35.6	(35.6-33.1)/5 = 0.4	0.4
Over fifteen	22.9	38.0	(38.0-33.1)/5 = 1.0	1.0
7. Most serious outstanding charge				
No outstanding charges	78.8	31.1	(31.1-33.1)/5 = -0.4	-0.4
Minor	10.6	43.2	(43.2-33.1)/5 = 2.0	2.0
Moderate	6.7	42.9	(42.9-33.1)/5 = 2.0	2.0
Serious	3.9	50.0	(50.0-33.1)/5 = 3.4	3.4
Major	0.0	0.0	(0.0-33.1)/5 = -6.6	-6.6
8. Severity of current offence				
Minor moderate	38.0	30.8	(30.8-33.1)/5 = -0.4	-0.4
Serious or major	60.9	34.6	(34.6-33.1)/5 = 0.4	0.4
Extreme	1.1	40.0	(40.0-33.1)/5 = 1.4	1.4
9. Sentence length				
1 day to 4 years	85.5	32.1	(32.1-33.1)/5 = -0.2	-0.2
5 to 9 years	12.3	38.6	(38.6-33.1)/5 = 1.2	1.2
10 to 24 years	1.7	50.0	(50.0-33.1)/5 = 3.4	3.4
Over 24 years	0.6	50.0	(50.0-33.1)/5 = 3.4	3.4
10. Street stability				
Above average	6.7	20.0	(20.0-33.1)/5 = -2.6	-2.6
Average	44.1	30.7	(30.7-33.1)/5 = -0.4	-0.4
Below average	49.2	39.3	(39.3-33.1)/5 = 1.2	1.2
11. Prior parole and / or statutory release - initial				
0	54.7	31.9	(31.9-33.1)/5 = -0.2	-0.2
1	18.4	36.3	(36.3-33.1)/5 = 0.6	0.6
2	3.9	43.8	(43.8-33.1)/5 = 2.2	2.2
3	1.1	100.0	(100.0-33.1)/5 = 13.4	13.4
4	6.7	33.3	(33.3-33.1)/5 = 0.0	0.0
5	3.4	18.8	(18.8-33.1)/5 = -2.8	-2.8
6	1.1	18.2	(18.2-33.1)/5 = -3.0	-3.0
7	1.7	100.0	(100.0-33.1)/5 = 13.4	13.4
8	1.1	20.0	(20.0-33.1)/5 = -2.6	-2.6

9	2.2	36.4	(36.4-33.1)/5 = 0.6	0.6
10	0.6	100.0	(100.0-33.1)/5 = 13.4	13.4
11	0.0	0.0	(0.0-33.1)/5 = -6.6	-6.6
12	0.6	50.0	(50.0-33.1)/5 = 3.4	3.4
13	1.7	50.0	(50.0-33.1)/5 = 3.4	3.4
15	0.0	0.0	(0.0-33.1)/5 = -6.6	-6.6
16	0.6	100.0	(100.0-33.1)/5 = 13.4	13.4
17	0.6	100.0	(100.0-33.1)/5 = 13.4	13.4
19	0.6	100.0	(100.0-33.1)/5 = 13.4	13.4
20	0.0	0.0	(0.0-33.1)/5 = -6.6	-6.6
21	0.6	100.0	(100.0-33.1)/5 = 13.4	13.4
24	0.6	100.0	(100.0-33.1)/5 = 13.4	13.4
26	0.6	100.0	(100.0-33.1)/5 = 13.4	13.4
40	0.0	0.0	(0.0-33.1)/5 = -6.6	-6.6
48	0.0	0.0	(0.0-33.1)/5 = -6.6	-6.6
11. Prior parole and / or statutory release – **revised**				
0	54.7	31.9	(31.9-33.1)/5 = -0.2	-0.2
>0	45.3	34.6	(34.6-33.1)/5 = 0.4	0.4
12. Age (at time of admission) - **initial**				
25 years of less	26.8	36.9	(36.9-33.1)/5 = 0.8	0.8
26	8.4	45.5	(45.5-33.1)/5 = 2.4	2.4
27	5.0	36.0	(36.0-33.1)/5 = 0.6	0.6
28	5.0	40.9	(40.9-33.1)/5 = 1.6	1.6
29	1.7	18.8	(18.8-33.1)/5 = -2.8	-2.8
30	6.1	47.8	(47.8-33.1)/5 = 3.0	3.0
31	2.8	29.4	(29.4-33.1)/5 = -0.8	-0.8
32	3.4	50.0	(50.0-33.1)/5 = 3.4	3.4
33	5.0	39.1	(39.1-33.1)/5 = 1.2	1.2
34	4.5	40.0	(40.0-33.1)/5 = 1.4	1.4
35 years or more	31.3	25.4	(25.4-33.1)/5 = -1.6	-1.6
12. Age (at time of admission) - **revised**				
35+ years of age	31.3	25.5	(25.5-33.1)/5 = -1.6	-1.6
Between 26 and 34 years	41.9	39.3	(39.3-33.1)/5 = 1.2	1.2
Less than 26	26.8	36.9	(36.9-33.1)/5 = 0.8	0.8

Appendix C: Development of the Burgess Scale Based on Incidents

CRS Items	Predictor Distributions (%)	Incident Rate per Predictor Category (%)	Intervals of a Full 5% Above or Below the Base Incident Rate (33.1%)	Final Score for the Category
1a. Involvement in prior incidents				
no	54.4	22.5	(22.5-27.2)/5 = -1.0	-1.0
yes	45.6	36.0	(36.0-27.2)/5 = 1.8	1.8
1b. Prior involvement in incidents in "greatest" or "high" severity category				
no	89.8	26.2	(26.2-27.2)/5 = -0.2	-0.2
yes	10.2	40.5	(40.5-27.2)/5 = 2.6	2.6
1c. Prior involvement during last five years of incarceration: • In an assault (no weapon or serious injury) • In a riot or major disturbance • In an assault (using a weapon or causing serious injury)				
no prior involvement during last 5 years of incarceration	78.9	25.3	(25.3-27.2)/5 = 2.6	-0.4
involvement in an assault (no weapon or serious injury)	17.7	34.2	(34.2-27.2)/5 = 2.6	1.4
either involvement in a riot or major disturbance or in an assault (using a weapon or causing serious injury)	3.4	71.4	(71.4-27.2)/5 = 2.6	8.8
1d. Involvement in one or more serious incidents prior to sentencing and / or pending placement from current commitment				
no	91.8	26.3	(26.3-27.2)/5 = -0.2	-0.2
yes	8.2	44.4	(44.4-27.2)/5 = 3.4	3.4
2. Escape history				
no escape or attempt	87.1	25.4	(25.4-27.2)/5 =-0.4	-0.4
an escape or attempt from minimum or community custody with no actual or threatened violence over two years ago	6.1	45.0	(45.0-27.2)/5 = 3.6	3.6
an escape or attempt from	5.4	66.7	(66.7-27.2)/5 = 7.8	7.8

minimum or community custody with no actual or threatened violence in last two years				
an escape or attempt from medium or maximum custody or an escape from minimum or community custody with actual or threatened violence over two years ago	0.0	0.0	(0.0-27.2)/5 = -5.4	-5.4
an escape or attempt from medium or maximum custody or an escape from minimum or community custody with actual or threatened violence in last two years OR two or more escapes from any level within the last five years	1.4	40.0	(40.0-27.2)/5 = 2.6	2.6
3. Street stability				
Above average	1.4	3.3	(3.3-27.2)/5 = -4.8	-4.8
Average	44.9	26.0	(26.0-27.2)/5 = -0.2	-0.2
Below average	53.7	35.0	(35.0-27.2)/5 = 1.6	1.6
4. Alcohol / drug use				
no identifiable problems	13.6	15.2	(15.2-27.2)/5 = -2.4	-2.4
abuse affecting one or more life areas	21.8	35.6	(35.6-27.2)/5 = 1.6	1.6
serious abuse affecting several life areas	64.6	29.8	(29.8-27.2)/5 = 1.6	0.6
5. Age (at time of sentencing) - **Initial**				
18 years or less	1.4	66.7	(66.7-27.2)/5 = 7.8	7.8
19	2.0	42.9	(42.9-27.2)/5 = 3.2	3.2
20	4.8	43.8	(43.8-27.2)/5 = 3.4	3.4
21	2.0	25.0	(25.0-27.2)/5 = -0.4	-0.4
22	1.4	14.3	14.3-27.2)/5 = -2.6	-2.6
23	6.8	33.3	(33.3-27.2)/5 = 1.2	1.2
24	1.4	11.8	(11.8-27.2)/5 = -3.0	-3.0
25	6.1	42.9	(42.9-27.2)/5 = 3.2	3.2

26	8.8	46.4	(46.4-27.2)/5 = 3.8	3.8
27	4.1	25.0	(25.0-27.2)/5 = -0.4	-0.4
28	4.1	26.1	(26.1-27.2)/5 = -0.2	-0.2
29	1.4	9.5	(9.5-27.2)/5 = -3.6	-3.6
30 years or more	55.8	25.2	(25.2-27.2)/5 = -0.4	-0.4
5. Age (at time of sentencing) - Revised				
30 years of age or more	55.8	25.2	(25.2-27.2)/5 = -0.4	-0.4
29 years of age or less	44.2	30.1	(30.1-27.2)/5 = 0.6	0.6
6. Number of prior convictions				
None	19.7	19.9	(19.9-27.2)/5 = -1.4	-1.4
One	8.8	22.0	(22.0-27.2)/5 = -1.0	-1.0
Two to four	19.0	27.2	(27.2-27.2)/5 = 0.0	0.0
Five to nine	17.7	32.5	(32.5-27.2)/5 = 1.0	1.0
Ten to fourteen	8.2	26.3	(26.3-27.2)/5 = -0.2	-0.2
Over fifteen	26.5	36.1	(36.1-27.2)/5 = 1.8	1.8
7. Most serious outstanding charge				
No outstanding charges	83.0	26.9	(26.9-27.2)/5 = 0.0	0.0
Minor	8.2	27.3	(27.3-27.2)/5 = 0.0	0.0
Moderate	4.8	25.0	(25.0-27.2)/5 = -0.4	-0.4
Serious	4.1	42.9	(42.9-27.2)/5 = 3.2	3.2
Major	0.00	0.00	(0.0-27.2)/5 = -5.4	-5.4
8. Severity of current offence				
Minor moderate	39.5	26.2	(26.2-27.2)/5 = -0.2	-0.2
Serious or major	59.2	27.6	(27.6-27.2)/5 = -5.4	0.0
Extreme	1.4	40.0	(40.0-27.2)/5 = -5.4	2.6
9. Sentence length				
1 day to 4 years	85.0	26.3	(26.3-27.2)/5 = -0.2	-0.2
5 to 9 years	12.9	33.3	(33.3-27.2)/5 = -1.2	1.2
10 to 24 years	0.7	16.7	(16.7-27.2)/5 = -2.2	-2.2
Over 24 years	1.4	100.0	(100.0-27.2)/5 = 14.6	14.6
10. Street stability				
Above average	2.0	5.0	(5.0-27.2)/5 = -4.4	-4.4
Average	44.9	25.7	(25.7-27.2)/5 = -0.2	-0.2
Below average	53.1	34.8	(34.8-27.2)/5 = 1.6	1.6
11. Prior parole and / or statutory release - initial				
0	57.1	27.4	(27.4-27.2)/5 = 0.0	0.0
1	17.7	28.6	(28.6-27.2)/5 = 0.2	0.2
2	4.1	37.5	(37.5-27.2)/5 = 2.0	2.0
3	0.7	50.0	(50.0-27.2)/5 = 4.6	4.6
4	6.1	25.0	(25.0-27.2)/5 = -0.4	-0.4
5	2.0	9.4	(9.4-27.2)/5 = -3.6	-3.6
6	2.0	27.3	(27.3-27.2)/5 = 0.0	0.0
7	0.0	0.0	(0.0-27.2)/5 = -5.4	-5.4
8	2.0	30.0	(30.0-27.2)/5 = 0.6	0.6

9	2.0	27.3	(27.3-27.2)/5 = 0.0	0.0
10	0.7	100.0	(100.0-27.2)/5 = 14.6	14.6
11	0.0	0.0	(0.0-27.2)/5 = -5.4	-5.4
12	1.4	100.0	(100.0-27.2)/5 = 14.6	14.6
13	1.4	33.3	(33.3-27.2)/5 = 1.2	1.2
15	0.0	0.0	(0.0-27.2)/5 = -5.4	-5.4
16	0.7	100.0	(100.0-27.2)/5 = 14.6	14.6
17	0.0	0.0	(0.0-27.2)/5 = -5.4	-5.4
19	0.0	0.0	(0.0-27.2)/5 = -5.4	-5.4
20	0.7	100.0	(100.0-27.2)/5 = 14.6	14.6
21	0.7	100.0	(100.0-27.2)/5 = 14.6	14.6
24	0.0	0.0	(0.0-27.2)/5 = -5.4	-5.4
26	0.7	100.0	(100.0-27.2)/5 = 14.6	14.6
40	0.0	0.0	(0.0-27.2)/5 = -5.4	-5.4
48	0.0	27.4	(27.4-27.2)/5 = 0.0	0.0
11. Prior parole and / or statutory release – revised				
0	57.1	27.4	(27.4-27.2)/5 = 0.0	0.0
>0	42.9	26.9	(26.9-27.2)/5 = 0.0	0.0
12. Age (at time of admission) – initial				
25 years of less	26.5	30.0	(30.0-27.2)/5 = 0.6	0.6
26	9.5	42.4	(42.4-27.2)/5 = 3.0	3.0
27	3.4	20.0	(20.0-27.2)/5 = -1.4	-1.4
28	5.4	36.4	(36.4-27.2)/5 = 1.8	1.8
29	0.0	0.0	(0.0-27.2)/5 = -5.4	-5.4
30	6.1	39.1	(39.1-27.2)/5 = 2.4	2.4
31	3.4	29.4	(29.4-27.2)/5 = 0.4	0.4
32	0.7	8.3	(8.3-27.2)/5 = -3.8	-3.8
33	2.7	17.4	(17.4-27.2)/5 = -2.0	-2.0
34	4.8	35.0	(35.0-27.2)/5 = 1.6	1.6
35 years or more	37.4	25.0	(25.0-27.2)/5 = -0.4	-0.4
12. Age (at time of admission) - revised				
35+ years of age	37.4	25.0	(25.0-27.2)/5 = -0.4	-0.4
Between 26 and 34 years	36.1	27.8	(27.8-27.2)/5 = 0.2	0.2
Less than 26	26.5	30.0	(30.0-27.2)/5 = 0.6	0.6

Appendix D: Example of ROC Analysis: Predicting Occurrence of Charges within Three Months of CRS Final Decision Date

Weighted CRS Score	Frequency	Probability Level	# of Correctly Predicted Events (True Positive)	# of Correctly Predicted Non-events (True Negative)	# of Nonevents Predicted as Events (False Positive)	# of Events Predicted as Nonevents (False Negative)	Sensitivity	1 - Specificity	Specificity	Mean of Sensitivity & Specificity	Percent in Maximum
23.2	1	0.77	1	237	0	100	0.01	0.00	1.00	0.50	0.3%
16.6	1	0.65	2	237	0	99	0.02	0.00	1.00	0.51	0.6%
16.2	1	0.65	3	237	0	98	0.03	0.00	1.00	0.51	0.9%
14.4	1	0.61	3	236	1	98	0.03	0.00	1.00	0.51	1.2%
14	1	0.60	3	235	2	98	0.03	0.01	0.99	0.51	1.5%
13.8	1	0.59	3	234	3	98	0.03	0.01	0.99	0.51	1.8%
13.2	1	0.58	3	233	4	98	0.03	0.02	0.98	0.51	2.1%
13	2	0.58	5	233	4	96	0.05	0.02	0.98	0.52	2.7%
12.6	1	0.57	5	232	5	96	0.05	0.02	0.98	0.51	3.0%
12.4	3	0.56	7	231	6	94	0.07	0.03	0.97	0.52	3.8%
12	1	0.55	8	231	6	93	0.08	0.03	0.97	0.53	4.1%
11.8	1	0.55	8	230	7	93	0.08	0.03	0.97	0.52	4.4%
11.6	2	0.54	9	229	8	92	0.09	0.03	0.97	0.53	5.0%
11.2	2	0.54	10	228	9	91	0.10	0.04	0.96	0.53	5.6%
11	1	0.53	11	228	9	90	0.11	0.04	0.96	0.54	5.9%
10.8	1	0.53	11	227	10	90	0.11	0.04	0.96	0.53	6.2%
10.6	4	0.52	12	224	13	89	0.12	0.05	0.95	0.53	7.4%
10.4	1	0.52	13	224	13	88	0.13	0.05	0.95	0.54	7.7%
10	1	0.51	13	223	14	88	0.13	0.06	0.94	0.53	8.0%
9.8	1	0.50	13	222	15	88	0.13	0.06	0.94	0.53	8.3%

9.6	1	0.50	14	222	15	87	0.14	0.06	0.94	0.54	8.6%
9.4	2	0.50	16	222	15	85	0.16	0.06	0.94	0.55	9.2%
9.2	2	0.49	16	220	17	85	0.16	0.07	0.93	0.54	9.8%
8.8	3	0.48	17	218	19	84	0.17	0.08	0.92	0.54	10.7%
8.6	2	0.48	18	217	20	83	0.18	0.08	0.92	0.55	11.2%
8.2	2	0.47	18	215	22	83	0.18	0.09	0.91	0.54	11.8%
8	3	0.46	19	213	24	82	0.19	0.10	0.90	0.54	12.7%
7.8	1	0.46	20	213	24	81	0.20	0.10	0.90	0.55	13.0%
7.6	2	0.45	22	213	24	79	0.22	0.10	0.90	0.56	13.6%
7.2	3	0.45	22	210	27	79	0.22	0.11	0.89	0.55	14.5%
7	2	0.44	23	209	28	78	0.23	0.12	0.88	0.55	15.1%
6.8	5	0.44	25	206	31	76	0.25	0.13	0.87	0.56	16.6%
6.6	2	0.43	27	206	31	74	0.27	0.13	0.87	0.57	17.2%
6.4	6	0.43	29	202	35	72	0.29	0.15	0.85	0.57	18.9%
6.2	3	0.42	30	200	37	71	0.30	0.16	0.84	0.57	19.8%
6	3	0.42	32	199	38	69	0.32	0.16	0.84	0.58	20.7%
5.6	1	0.41	32	198	39	69	0.32	0.16	0.84	0.58	21.0%
5.4	1	0.41	32	197	40	69	0.32	0.17	0.83	0.57	21.3%
5.2	1	0.40	33	197	40	68	0.33	0.17	0.83	0.58	21.6%
4.6	1	0.39	34	197	40	67	0.34	0.17	0.83	0.58	21.9%
4.4	3	0.38	35	195	42	66	0.35	0.18	0.82	0.58	22.8%
4.2	3	0.38	36	193	44	65	0.36	0.19	0.81	0.59	23.7%
3.8	5	0.37	37	189	48	64	0.37	0.20	0.80	0.58	25.1%
3.6	5	0.37	40	187	50	61	0.40	0.21	0.79	0.59	26.6%
3.4	6	0.36	45	186	51	56	0.45	0.22	0.78	0.62	28.4%
3.2	5	0.36	45	181	56	56	0.45	0.24	0.76	0.60	29.9%
3	6	0.35	47	177	60	54	0.47	0.25	0.75	0.61	31.7%
2.8	8	0.35	50	172	65	51	0.50	0.27	0.73	0.61	34.0%
2.6	2	0.35	51	171	66	50	0.50	0.28	0.72	0.61	34.6%

2.4	2	0.34	53	171	66	48	0.52	0.28	0.72	0.62	35.2%
2.2	2	0.34	53	169	68	48	0.52	0.29	0.71	0.62	35.8%
2	2	0.33	54	168	69	47	0.53	0.29	0.71	0.62	36.4%
1.8	2	0.33	55	167	70	46	0.54	0.30	0.70	0.62	37.0%
1.6	2	0.33	56	166	71	45	0.55	0.30	0.70	0.63	37.6%
1.4	1	0.32	56	165	72	45	0.55	0.30	0.70	0.63	37.9%
1.2	2	0.32	57	164	73	44	0.56	0.31	0.69	0.63	38.5%
1	5	0.31	60	162	75	41	0.59	0.32	0.68	0.64	39.9%
0.8	2	0.31	61	161	76	40	0.60	0.32	0.68	0.64	40.5%
0.6	9	0.31	63	154	83	38	0.62	0.35	0.65	0.64	43.2%
0.4	8	0.30	63	146	91	38	0.62	0.38	0.62	0.62	45.6%
0	3	0.29	64	144	93	37	0.63	0.39	0.61	0.62	46.4%
-0.2	3	0.29	64	141	96	37	0.63	0.41	0.59	0.61	47.3%
-0.4	4	0.29	66	139	98	35	0.65	0.41	0.59	0.62	48.5%
-0.6	2	0.28	66	137	100	35	0.65	0.42	0.58	0.62	49.1%
-0.8	2	0.28	66	135	102	35	0.65	0.43	0.57	0.61	49.7%
-1	8	0.28	69	130	107	32	0.68	0.45	0.55	0.62	52.1%
-1.2	4	0.27	70	127	110	31	0.69	0.46	0.54	0.61	53.3%
-1.4	2	0.27	70	125	112	31	0.69	0.47	0.53	0.61	53.8%
-1.6	5	0.27	72	122	115	29	0.71	0.49	0.51	0.61	55.3%
-1.8	2	0.26	73	121	116	28	0.72	0.49	0.51	0.62	55.9%
-2	5	0.26	73	116	121	28	0.72	0.51	0.49	0.61	57.4%
-2.2	3	0.25	74	114	123	27	0.73	0.52	0.48	0.61	58.3%
-2.4	7	0.25	77	110	127	24	0.76	0.54	0.46	0.61	60.4%
-2.6	2	0.25	78	109	128	23	0.77	0.54	0.46	0.62	60.9%
-2.8	8	0.24	82	105	132	19	0.81	0.56	0.44	0.63	63.3%
-3	2	0.24	83	104	133	18	0.82	0.56	0.44	0.63	63.9%
-3.2	3	0.24	85	103	134	16	0.84	0.57	0.43	0.64	64.8%
-3.4	7	0.23	85	96	141	16	0.84	0.59	0.41	0.62	66.9%

-3.6	3	0.23	85	93	144	16	0.84	0.61	0.39	0.62	67.8%
-3.8	3	0.23	86	91	146	15	0.85	0.62	0.38	0.62	68.6%
-4	1	0.22	86	90	147	15	0.85	0.62	0.38	0.62	68.9%
-4.2	7	0.22	89	86	151	12	0.88	0.64	0.36	0.62	71.0%
-4.4	2	0.22	89	84	153	12	0.88	0.65	0.35	0.62	71.6%
-4.6	4	0.22	90	81	156	11	0.89	0.66	0.34	0.62	72.8%
-4.8	3	0.21	90	78	159	11	0.89	0.67	0.33	0.61	73.7%
-5	2	0.21	90	76	161	11	0.89	0.68	0.32	0.61	74.3%
-5.2	1	0.21	90	75	162	11	0.89	0.68	0.32	0.60	74.6%
-5.4	5	0.20	91	71	166	10	0.90	0.70	0.30	0.60	76.0%
-5.6	2	0.20	91	69	168	10	0.90	0.71	0.29	0.60	76.6%
-5.8	1	0.20	91	68	169	10	0.90	0.71	0.29	0.59	76.9%
-6	5	0.19	92	64	173	9	0.91	0.73	0.27	0.59	78.4%
-6.2	3	0.19	93	62	175	8	0.92	0.74	0.26	0.59	79.3%
-6.4	5	0.19	93	57	180	8	0.92	0.76	0.24	0.58	80.8%
-6.8	3	0.18	94	55	182	7	0.93	0.77	0.23	0.58	81.7%
-7	1	0.18	94	54	183	7	0.93	0.77	0.23	0.58	82.0%
-7.2	1	0.18	95	54	183	6	0.94	0.77	0.23	0.58	82.2%
-7.4	5	0.18	96	50	187	5	0.95	0.79	0.21	0.58	83.7%
-7.8	1	0.17	96	49	188	5	0.95	0.79	0.21	0.58	84.0%
-8.2	2	0.17	96	47	190	5	0.95	0.80	0.20	0.57	84.6%
-8.4	2	0.16	96	45	192	5	0.95	0.81	0.19	0.57	85.2%
-8.6	1	0.16	96	44	193	5	0.95	0.81	0.19	0.57	85.5%
-8.8	3	0.16	97	42	195	4	0.96	0.82	0.18	0.57	86.4%
-9	1	0.16	98	42	195	3	0.97	0.82	0.18	0.57	86.7%
-9.2	1	0.15	98	41	196	3	0.97	0.83	0.17	0.57	87.0%
-9.6	5	0.15	99	37	200	2	0.98	0.84	0.16	0.57	88.5%
-10	2	0.14	99	35	202	2	0.98	0.85	0.15	0.56	89.1%
-10.2	6	0.14	99	29	208	2	0.98	0.88	0.12	0.55	90.8%

-10.6	3	0.14	100	27	210	1	0.99	0.89	0.11	0.55	91.7%
-10.8	1	0.14	100	26	211	1	0.99	0.89	0.11	0.55	92.0%
-11	3	0.13	100	23	214	1	0.99	0.90	0.10	0.54	92.9%
-11.4	3	0.13	100	20	217	1	0.99	0.92	0.08	0.54	93.8%
-11.6	1	0.13	100	19	218	1	0.99	0.92	0.08	0.54	94.1%
-11.8	1	0.12	100	18	219	1	0.99	0.92	0.08	0.53	94.4%
-12.6	2	0.12	100	16	221	1	0.99	0.93	0.07	0.53	95.0%
-12.8	1	0.12	100	15	222	1	0.99	0.94	0.06	0.53	95.3%
-14.2	1	0.10	100	14	223	1	0.99	0.94	0.06	0.52	95.6%
-14.4	1	0.10	100	13	224	1	0.99	0.95	0.05	0.52	95.9%
-14.6	1	0.10	101	13	224	0	1.00	0.95	0.05	0.53	96.2%
-15	1	0.10	101	12	225	0	1.00	0.95	0.05	0.53	96.4%
-15.6	5	0.09	101	7	230	0	1.00	0.97	0.03	0.51	97.9%
-16	2	0.09	101	5	232	0	1.00	0.98	0.02	0.51	98.5%
-16.4	5	0.09	101	0	237	0	1.00	1.00	0.00	0.50	100.0%

Appendix E: Inter-item (Spearman) Correlations of Reweighted CRS

	M (SD)	1	2	3	4	5	6	7	8	9	10	11	12
1. Incident involvement	-0.08 (3.31)	1											
2. Escape history	-0.06 (0.81)	0.24**	1										
3. Street stability	-0.01 (1.54)	0.33**	0.12	1									
4. Alcohol / drug use	0.06 (1.86)	0.31**	0.10	0.58**	1								
5. Age (at time of sentencing)	-0.04 (0.69)	0.08	-0.2	0.09	0.10	1							
6. Number of prior convictions	0.04 (0.95)	0.29**	0.17**	0.40**	0.44**	-0.14**	1						
7. Most serious outstanding charge	0.01 (1.02)	0.09	0.10	0.02	0.11	0.03	0.08	1					
8. Severity of current offence	0.08 (0.41)	0.14*	-0.02	0.13	0.07	0.11	-0.07	-0.03	1				
9. Sentence length	0.00 (0.60)	0.04	0.06	0.00	-0.08	-0.02	-0.07	0.01	0.20**	1			
10. Street stability	0.02 (1.19)	0.33**	0.13	0.93**	0.56**	0.07	0.40**	0.00	0.14*	0.02	1		
11. Prior parole and/or statutory release	0.06 (0.30)	0.40**	0.18***	0.26**	0.31**	-0.10	0.38**	0.04	0.06	-0.07	0.29**	1	
12. Age (at time of admission)	-0.03 (1.31)	0.15*	0.06	0.09	0.13	0.44**	-0.02	0.07	0.13	0.06	0.09	0.06	1

*$p < 0.05$ ** $p < 0.001$, *** $p < 0.0001$

Appendix F: Consistency between the Reweighted CRS and Other Security Level Decisions

Table 1

Comparison between CRS Recommendations (Original and Reweighted)

Overall Percentages		Based on Reweighted CRS		
		Minimum	Medium	Maximum
Based on Original CRS	Minimum	**75.94**	25.93	0.00
	Medium	24.06	**70.78**	43.75
	Maximum	0.00	3.29	**56.25**
Aboriginal Percentages				
		Minimum	Medium	Maximum
Based on Original CRS	Minimum	**58.97**	14.13	0.00
	Medium	41.03	**83.70**	33.33
	Maximum	0.00	2.17	**66.67**
Non-Aboriginal Percentages				
		Minimum	Medium	Maximum
Based on Original CRS	Minimum	**78.85**	33.11	0.00
	Medium	21.15	**62.91**	50.00
	Maximum	0.00	3.97	**50.00**

Table 2

Comparison between Final Decision and Reweighted CRS Recommendations

Overall Percentages		Based on Reweighted CRS		
		Minimum	Medium	Maximum
Based on Final Decision	**Minimum**	**73.31**	3.00	0.00
	Medium	26.32	**72.43**	59.38
	Maximum	0.38	23.46	**40.63**
Aboriginal Percentages				
		Minimum	Medium	Maximum
Based on Final Decision	**Minimum**	**43.59**	15.22	0.00
	Medium	53.85	**79.35**	58.33
	Maximum	2.56	5.43	**41.67**
Non-Aboriginal Percentages				
		Minimum	Medium	Maximum
Based on Final Decision	**Minimum**	**78.41**	28.48	0.00
	Medium	21.59	**68.21**	60.00
	Maximum	0.00	3.31	**40.00**