

# ASSESSING DYNAMIC RISK AND DYNAMIC STRENGTH CHANGE PATTERNS AND THE RELATIONSHIP TO REOFFENDING AMONG WOMEN ON COMMUNITY SUPERVISION

KAYLA A. WANAMAKER 

SHELLEY L. BROWN 

*Carleton University*

---

This study examines how dynamic risk and strength factors change over time and whether these changes are predictive of reoffending outcomes. The sample includes 2,877 Canadian women under community supervision with Service Planning Instrument reassessment data. Over a 30-month period, patterns of change in total dynamic risk and strength scores were examined. Change parameters were entered into a series of logistic regression models, linking change to three reoffending outcomes: technical violations, any new charges, and new violent charges. Overall, total dynamic risk scores decreased, and total dynamic strength scores increased over time. Change in total dynamic risk scores predicted any new charges and technical violations, whereas change in total dynamic strength scores only predicted technical violations. Findings demonstrated the utility of reassessing dynamic risk and strength scores over time and support the incorporation of strengths-based approaches with women involved in the criminal justice system.

**Keywords:** risk assessment; women; dynamic risk; strength; community supervision; reoffending

---

Risk assessment is critical for informing case management plans for those involved in the criminal justice system, which includes determining appropriate supervision intensity and programming targets. Although an abundance of research has been conducted focusing on risk assessment with justice-involved men, including more recent studies focusing on reassessment and patterns of change (e.g., Lloyd et al., 2020), research focusing on risk assessment with justice-involved women is in its infancy. Moreover, the research that exists with justice-involved women have focused predominately on whether traditional, gender-neutral risk assessment tools—that is, tools that use a one-size-fits-all approach to assess risk, regardless of one’s gender—adequately predict reoffending outcomes for women (e.g., Scanlan et al., 2020). In contrast, a small subset of risk assessment research

---

**AUTHORS’ NOTE:** *This article is based on Kayla Wanamaker’s doctoral dissertation (2020) titled, “A Multi-Wave Longitudinal Examination of How Strengths and Risks Inform Risk Assessment and Treatment Profiles for Justice-Involved Men and Women Using the Service Planning Instrument (SPIn).” The authors would like to thank Orbis Partners Inc. and Alberta Justice and Solicitor General for providing access to SPIn and reoffense data sets. Correspondence concerning this article should be addressed to Kayla A. Wanamaker, Department of Psychology, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, Canada K1S 5B6; e-mail: kaylawanamaker@cmail.carleton.ca.*

CRIMINAL JUSTICE AND BEHAVIOR, 202X, Vol. XX, No. X, Month 2021, 1–21.

DOI: 10.1177/00938548211026706



Article reuse guidelines: [sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

© 2021 International Association for Correctional and Forensic Psychology

has focused on tools developed specifically for women that include risk factors thought to be women-specific, such as self-esteem and self-efficacy, mental health issues, intimate partner violence, and childhood abuse (e.g., Women's Risk Needs Assessment; Van Voorhis et al., 2010). Regardless, studies examining the utility of assessment with women, both gender-neutral and women-specific tools, have predominately used single-wave designs. One exception is Greiner et al. (2015) who used a multiwave design to examine the use of risk assessment with women overtime and whether change in scores predict reoffending outcomes. Given the lack of research assessing change among women, the purpose of this study is to examine how strength and risk factors change over time among women on community supervision and how change in scores predict reoffending outcomes.

### RISK ASSESSMENT

Briefly, risk assessment tools have traditionally focused on examining factors that are predictive of criminal behavior, typically classified as either static risk factors or dynamic risk factors. Static risk factors, such as "age at first arrest," are historical factors that are unchangeable as a function of intervention (Zamble & Quinsey, 1997). Because these factors do not change, they cannot act as treatment targets. In comparison, dynamic risk factors, such as "criminogenic attitudes," are changeable factors and therefore can be modified by correctional treatment (Bonta & Andrews, 2017). These factors, used interchangeably with the term *criminogenic needs*, measure an offender's propensity to commit an offense at a specific time and inform rehabilitation. An abundance of research has found support for the use of assessment tools that combine static and dynamic risk variables, both practically and in terms of predictability (Bonta & Andrews, 2017). Research has also found that risk to reoffend does change over time and that reassessment of dynamic factors improves prediction of outcomes (Lloyd et al., 2020). However, understanding how risk changes over time is limited.

Empirical findings have also demonstrated that strengths can inform rehabilitation efforts (e.g., Brown et al., 2020). Within the context of criminal justice research, the concept of strength has been used in different capacities. Some scholars have suggested that a strength is merely the absence of a risk factor, proposing that a variable cannot be viewed as both a risk and a strength (Ogloff & Davis, 2004). However, there is evidence to suggest that strengths can occur simultaneously with risk factors and can account for the variability in specific outcomes (e.g., Jones et al., 2015). Other researchers further contend that a strength must add something positive to an individual's life (Lodewijks et al., 2010). For the purpose of this study, strengths are defined as positive internal resources (e.g., individual's priorities, goals, and values) or external resources (e.g., prosocial peers and supportive family) that an individual has available to them (Laws & Ward, 2011). Despite some positive findings associated with the use of strengths in risk assessment (e.g., Jones et al., 2015), there remains considerable debate in the literature around the operationalization, measurement, and utility of strengths (see Wanamaker et al., 2018, for a detailed discussion of incorporating strengths in risk assessment protocols). Further strengths-based research is required.

### RISK ASSESSMENT RESEARCH WITH WOMEN

Risk assessment research focusing on justice-involved women is relatively limited in comparison with research on justice-involved men. The majority of research with women

tend to focus on pathways into the criminal justice system (e.g., Belknap, 2015) or concentrate on identifying variables that should be incorporated in risk assessment measures. Researchers in this area argue that existing risk/needs assessment models were developed on male samples and fail to include factors that are most relevant for women (Van Voorhis et al., 2010). As such, women-focused assessment models have suggested that there are several variables especially salient for women, which include parental stress, family support, self-efficacy, educational assets, housing safety, anger/hostility, current mental health and relationship dysfunction, and victimization (Van Voorhis et al., 2010; Wardrop et al., 2019). This body of research has also focused on identifying appropriate classification cutoff scores for women, positing that cutoff scores for men may be inappropriate for use with women, leading to either over- or underclassification (Blanchette & Brown, 2006). This body of work, however, has predominately utilized single-wave designs, many of which do not incorporate rigorous statistical approaches or examine how women's scores on various risk and strength factors change over time.

One assessment tool that is considered gender-informed, in that it considers factors that are gender-neutral and gender-salient, is the Service Planning Instrument (SPIn; Orbis Partners, 2003). The SPIn is a risk, need, and strength assessment and case management planning instrument used with adults who are incarcerated or in community-based justice settings. It is comprised of 90 static and dynamic risk and/or strength items found to be predictive of reoffending for both men and women and are divided into 11 content domains: criminal history, response to supervision (e.g., institutional misconducts, violations), aggression, substance use, social influences (e.g., negative and prosocial peers), family, employment and education, attitudes, social and cognitive skills, stability (e.g., living situation), and mental health.

A recent study conducted by Brown et al. (2020) used logistic regression and area under the receiver operator characteristic (ROC) curves (AUCs) to examine the predictability of SPIn Full Assessment total strength and total risk scores. Results were disaggregated by two mixed-gender samples: one Clark County, Washington probation sample ( $N = 1,975$ ) and one community sample from Alberta, Canada ( $N = 20,537$ ).<sup>1</sup> For the Washington sample, both the risk scores ( $AUC = .64$ , 95% confidence interval [CI] = [.61, .66]) and strength scores ( $AUC = .64$ , 95% CI = [.61, .66]) were moderately predictive of any new offenses (charges and/or convictions) within a 2-year fixed follow-up. For the Canadian sample, both risk scores ( $AUC = .59$ , 95% CI = [.58, .60]) and strength scores ( $AUC = .58$ , 95% CI = [.57, .59]) predicted any reconviction over a 3-year fixed follow-up, albeit the effect size was small. Longitudinal reassessment data are needed to clarify the relationship between dynamic risks, strengths, and reoffending outcomes among women and how changes in risk and strength variables may further enhance predictive accuracy.

### MULTIWAVE ASSESSMENT OF DYNAMIC RISKS AND STRENGTHS

While many studies have incorporated the use of multiwave designs to assess change in risk assessment scores over time, the majority have focused on men (e.g., Babchishin & Hanson, 2020) and tend to use small samples or two-way designs (e.g., Lewis et al., 2013). Research employing two-way designs can meaningfully examine change prior to and post intervention; however, when assessing change over a longer time period without examining

the effects of an intervention, it has been suggested that studies include at least three time-points to increase the accuracy of detecting change (Brown et al., 2009). It is particularly important to incorporate more assessment points when the nature and/or frequency of intervention(s) is unknown.

In studies that have used three or more timepoints, change has been assessed using various Cox regression survival analysis models, which compare characteristics of individuals at the time of reoffense with other individuals still at risk. Typically, these models have assessed incremental changes in prediction over initial scores, that is, examining improvements to prediction at each assessment timepoint or improved model fit of the most recent assessment (proximal) compared to the initial assessment (distal). Other change studies have examined whether averaging reassessment scores and incorporating them into Cox regression models add incrementally to the prediction of reoffending outcomes over initial assessment scores alone.

Greiner et al. (2015) is the only study that has assessed the relationship between change in dynamic risk scores and reoffending specifically for women. This study examined how dynamic factors change across four timepoints, collected at 6-month intervals, and how dynamic reassessments improve predictive accuracy. Using a sample of 497 justice-involved Canadian women released into the community from federal custody, they found that several dynamic risk factors measured using the Community Intervention Scale (Dowden et al., 2001), including employment, marital/family, community functioning, personal/emotional, criminal associates, and criminal attitudes, significantly decreased among offenders who refrained from reoffending. However, Cox regression with time-dependent covariates and AUC analyses revealed that while all seven dynamic risk predictors were significantly related to survival time, the best fit model at Time 1 included two variables—associates and attitudes (AUC = .66). In contrast, the best fit model at the last assessment before failure also included two variables—associates and employment needs (AUC = .70). Notably, the last assessment before failure was more predictive than Time 1 assessment information.

There have been two multiwave studies examining the reassessment of dynamic risks and strengths: Hanby (2013) and Lloyd et al. (2020). These studies used reassessment data over a 12-month period from the Dynamic Risk Assessment of Offender Re-entry (DRAOR; Serin, 2007) using 3,498 parolees from New Zealand (246 women, 3,249 men, although results were not disaggregated by sex). Hanby (2013) used hierarchical linear modeling (HLM) and Cox regression survival analysis to examine monthly DRAOR reassessments and found that dynamic risk variables decreased, whereas protective factors increased over time. However, for those who were not successful, in the month prior to reoffending, strength scores dropped. Findings from the Cox regression analyses indicated that later assessments had lower predictive accuracy than earlier assessments. Monthly average protective factor scores predicted reconvictions for the first 4 months of parole only. Although Hanby (2013) examined how scores at each timepoint predict reconvictions, this study did not assess the predictive nature of change scores.

Lloyd et al. (2020) compared three Cox regression models. The first model included only static risk scores from the Risk of Reconviction  $\times$  Risk of Reimprisonment scale (RoC  $\times$  RoI; Bakker et al., 1998). The second model included static scores and initial stable, acute, and protective domain scores. The third model included static scores and change scores for the stable, acute, and protective domains. Results demonstrate that Model 2 added

incrementally to the prediction of static risk scores alone,  $\chi^2(3) = 158.85, p < .001$ . However, Model 3 added incrementally to the prediction of static risk scores to a higher degree,  $\chi^2(6) = 300.91, p < .001$ . Although this demonstrates the importance of reassessing strengths, this study focused on men and used the DRAOR to assess strengths.

Although research focusing on the dynamic nature of risk factors among women is limited, the results from Greiner et al. (2015) coupled with research assessing dynamic factors with samples of men (e.g., Brown et al., 2009) demonstrate that the incorporation of dynamic risk factors, in combination with static risk factors, improves the prediction of recidivism. Limited research has examined the dynamic nature of strength factors. Each of these studies used Cox regression survival analysis to examine specific timepoints most predictive of reoffending outcomes; however, these studies did not examine whether the actual change in scores was predictive of reoffending outcomes.

## PURPOSE

The purpose of this study is to examine the dynamic nature of SPIn-assessed dynamic risk and strength scores. To date, no studies have assessed how changes in SPIn-assessed total dynamic risk and strength scores predict recidivism among justice-involved women. This study uses multiwave reassessment data from women on community supervision to assess: (a) patterns of change in total dynamic risk and strength scores, and (b) the relationship between changes in total dynamic risk and strength scores and reoffending outcomes.

## METHOD

### PARTICIPANTS

The sample comprised women from Alberta, Canada,<sup>2</sup> serving a provincial sentence<sup>3</sup> that either started immediately under community supervision or post-release from a provincial correctional facility. The sample included 2,877 women who started community supervision between 2009 and 2012 and who had between three and five completed SPIn Full Assessments ( $M = 3.8, SD = 0.86, \text{median} = 3.0$ ). There was a fixed 3-year follow-up period from the time of the initial assessment. As part of the inclusion criteria, all initial assessments occurred within 90 days of the start of supervision to ensure information was reflective of true initial scores and to ensure change parameters accurately represented individuals' change trajectory.

### MEASURES

#### The SPIn—Total Dynamic Risk and Strength Scores

The SPIn was scored by supervising probation officers. Information was obtained from semi-structured interviews administered by the probation officers along with file reviews. Probation officers are required to complete a total of 4 days of training—2 days of training on how to administer the SPIn, and another 2 days of training on how to apply the SPIn results to case planning and case management. The SPIn includes a Prescreen and a Full Assessment component. Of the 90 total items, 35 are used to calculate Prescreen risk and strength scores. While the Prescreen strength score is rarely used in practice, the Prescreen

risk score is used for making initial supervision classification–level decisions. This study uses the total dynamic risk and total dynamic strength score from the SPIn Full Assessment.<sup>4</sup>

To calculate the total dynamic risk score, eight dynamic risk scores were aggregated. These dynamic domains scores were from the family, employment, substance use, aggression, social, attitudes, social/cognitive skills, and stability domains and scores ranged from 0 to 75+ (scores over 75 are considered very high). The total dynamic strength score<sup>5</sup> was created by aggregating seven dynamic strength scores from the family, employment, social, aggression, attitudes, cognitive/social skills, and stability domains. The total dynamic strength scores ranged from 0 to 57+ (scores over 57 are considered very high). This study found good internal consistency for the total dynamic risk domain ( $\alpha = .76$ ) and dynamic strength domain ( $\alpha = .85$ ).

Many items creating the dynamic strength and dynamic risk scores are scored via a pooled 5-point Likert-type fashion scale such that they are scored as strengths or risks. Specifically, risk and strength represent opposite poles of the construct being assessed—take, for example, current intimate relationships. This item can be rated on the risk side as either a 2 (*high degree of instability and conflict, offender expresses high dissatisfaction*) or a 1 (*some conflict and dissatisfaction evident in the relationship*). Conversely, it can be rated on the strength side as a 2 (*high degree of stability, satisfaction, and commitment to the relationship*) or a 1 (*stability of the relationship evident, offender expresses satisfaction*). Finally, the item can instead be rated as neutral (score of 0), which indicates that the item is neither a strength nor a risk factor (*minimal satisfaction in relationship or no current marital relationship*). Therefore, while many domains contain both strength and risk scores, these strength and risk scores are often rated using the same items. The remaining items are not scored in a pooled fashion. Instead, these items are scored as either pure risk factors or pure strength factors (i.e., either dichotomous yes/no items, or “check all that apply” items, such as “gang association” or “partner has prosocial influence”). For additional information on the SPIn, see Jones and Robinson (2018).

The SPIn is used throughout Alberta and policy indicates that initial assessments are conducted within 45 days of the start of court-ordered community supervision or release from custody, and reassessments are completed every 6 months. The SPIn is not required under four conditions. First, if the supervision period is 3 months or less. Second, if the court order does not have a reporting condition. Third, if the client is convicted of a new charge while under supervision and there is a recently completed assessment. And finally, if within 45 days of the start of supervision, the client leaves the province of Alberta.

### Covariates

Three covariates were included in the current analyses: total static risk score, age, and Indigenous status. Total static risk score was made up of static risk items from seven content domains from the SPIn Full Assessment, including family, employment, criminal history, response to supervision, aggression, substance use, and stability. The average total static risk score was 20.62 ( $SD = 15.99$ ; range = 0–112). Age, which was included to account for any differences between younger and older women, refers to age at the time of initial assessment. In the current sample, women’s ages ranged from 16 to 77 years old. Finally, given the high proportion of Indigenous women involved in the criminal justice system in Canada, a dichotomous variable, Indigenous status, was used to identify Indigenous and

non-Indigenous women. Indigenous status refers to those who self-identify as First Nations, Inuit, or Metis.

### **Criminal Outcome**

Three dichotomous (yes/no) measures based on reoffense records indicating recontact with Alberta correctional services were used. The first outcome was any new charge(s), which included any new charges that are nonviolent, sexual, or violent in nature, but excludes any technical violations. The second outcome was any new violent charge(s), which included uttering threats, all forms of assault (including causing bodily harm, assault with a weapon, assault of a peace officer, and simple assaults), any weapon-related offenses (including pointing a firearm, possession, and careless storage), harassment, robbery, dangerous driving or operation causing bodily harm, damage by arson, and any murder charges. Importantly, violent charges exclude any technical violation, any new nonviolent charge, and any new sexual charge. The third outcome was technical violation(s) that included any breaches of court-ordered or community supervision conditions resulting in a failure to comply or failure to appear. Outcomes were assessed over a 3-year fixed follow-up from initial SPIn assessment.

### **ANALYTIC APPROACH**

The first step to assess change over time is to examine whether the SPIn meets the measurement invariance assumption. A violation of measurement invariance indicates that a measure may not accurately assess true change in scores over time, but rather change in scores may be due to measurement error.

### **Patterns of Change**

HLM was used to examine patterns of change at the person level and the extent to which covariates were related to different patterns of change. Three models were created for both the total dynamic risk and strength scores from the SPIn Full Assessment. The unconditional means model describes the variation in the initial scores and does not consider change. The unconditional growth model includes time as the only predictor to examine within-individual effects (e.g., Level 1). The conditional model includes time and Level 2 predictors to examine both within- and between-individual effects. Time can be entered into the model as a linear, quadratic, or cubic trend, determined by the loglikelihood ratio test results. Time was centered around the initial assessment date in months (i.e., initial assessment is considered timepoint 0, each subsequent month that passes is considered a new timepoint).

### **Prediction**

To examine how the previously assessed patterns of change in dynamic needs and dynamic strengths relate to criminal outcomes, a two-stage HLM was used, which is an extension of basic HLM. Analyses included taking the individual intercepts and slopes from the HLM analyses and inputting them into a regression equation. Parameter estimates from the conditional model that included between-individual effects (i.e., covariates) were retained. See Yang et al. (2017) for more details on the two-stage HLM approach.

## RESULTS

### SAMPLE DESCRIPTIVES

Overall, the average age of the sample was 32.9 years old ( $SD = 10.8$ ) and over a quarter of the sample were Indigenous women (28.8%). Based on SPIn Full Assessment static risk scores, over half of the sample were low risk (57.6%), a third were moderate risk (34.7%), and very few were high risk (6.9%). For the total dynamic risk score, the mean was 18.7 ( $SD = 15.9$ ), and for the total dynamic strength score, the mean was 26.9 ( $SD = 16.5$ ). Roughly two thirds of the sample had at least one mental health concern (64.8%) that includes but is not limited to anxiety, a major mental disorder, suicidal ideation, and homicidal ideation. Just below a fifth of the sample (18.6%) had served at least one custodial sentence. For index offenses, 37.6% committed a nonviolent offense, 29.8% committed a violent offense, 6.8% committed an administration of justice offense, and 1.3% committed a sexual offense. For recidivism, 10.1% were charged with a new violent offense, 24.9% were charged with any new offense, and 19.7% committed a technical violation within 3 years post initial SPIn Assessment.

### DATA SCREENING

Prior to running analyses, the data were screened for missing values, outliers, normality, linearity, and multicollinearity. Missing assessment information was a function of time; data across timepoints were missing because it was not available due to length of follow-up, with less than 3% of cases missing data at Timepoints 1 to 3, 57.3% ( $n = 1,648$ ) missing data at Timepoint 4, and 80% ( $n = 2,301$ ) missing data at Timepoint 5. Thus, the data are considered to be missing at random (MAR). A sensitivity analysis indicated that outliers (defined as extreme standardized residuals in the 0 and 99th percentile) did not influence overall results, and results were presented with all cases included.

### MEASUREMENT INVARIANCE

Two types of measurement invariance models were tested: the configural model and the scalar model. The configural model is a baseline model used to compare to more restrictive models. The scalar invariance model is a more restrictive model that assesses whether factor loadings and intercepts are equivalent across timepoints (Putnick & Bornstein, 2016). Essentially, achieving scalar measurement invariance provides evidence to support the utility of mean comparisons across timepoints. Upon testing measurement invariance, results indicated slight improvement when comparing the scalar with the configural model for both dynamic risk (Tucker–Lewis index [ $TLI_{\text{change}}$ ] = .001, root mean square error of approximation [ $RMSEA_{\text{change}}$ ] = .003) and dynamic strength ( $RMSEA_{\text{change}}$  = .002), suggesting that the scalar model fits the data better (see Table 1). In other words, the SPIn total dynamic strength and total dynamic risk scores are assessing the same construct over time.

### PATTERNS OF CHANGE RESULTS

#### Dynamic Risk

The initial assessment SPIn total dynamic risk scores ranged from 0 to 163 ( $M = 18.73$ ,  $SD = 15.89$ ). Subsequent SPIn assessments occurred at intervals spanning 3 to 30 months post initial assessment with scores ranging from 0 to 147 ( $M = 18.77$ ,  $SD = 15.39$ ).

**TABLE 1: Model Fit Indices and Nested Model Comparisons for Testing Invariance of the SPIn**

Dynamic strength	CFI	TLI	RMSEA [95% CI]
Configural <sup>a</sup>	.997	.996	.056 [.078, .081]
Scalar <sup>b</sup>	.997	.997	.053 [.050, .055]
Dynamic risk	CFI	TLI	RMSEA [95% CI]
Configural	.997	.997	.050 [.047, .052]
Scalar	.997	.997	.048 [.045, .050]

Note. SPIn = Service Planning Instrument; CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation; CI = confidence interval.

<sup>a</sup>Configural = No equality constraints. <sup>b</sup>Scalar = Factor loadings and intercepts constrained to be equal across all timepoints.

The loglikelihood value associated with the random effects linear function (73,447.3) was smaller than the quadratic function (74,983.7), and as such, the likelihood ratio test could not be ran indicating that the linear random effects trend best fit the data. In addition, the intraclass correlation coefficient (ICC) was calculated by dividing the intercept variance by the total variance (intercept and residual variance), see model results presented in Table 2. The ICC was .852 (201.93 / [35.01 + 201.93]), which indicates that 85.2% of the variance in initial total dynamic risk scores can be explained by differences among justice-involved women.

*Unconditional growth model: Level 1 change over time.* As seen in Table 2, results showed that, on average, the initial total dynamic risk score was 18.54 and the average rate of change was a decrease in scores by 0.04 points per month while on community supervision. The variance in initial status ( $\tau_{00} = 231.54$ ) and rate of change ( $\tau_{11} = 0.29$ ) were both significantly different from 0 ( $p < .001$ ), indicating that there was significant variability in both initial total dynamic risk scores and in the rate of change on these scores. Furthermore, the covariance between initial score and change over time ( $\tau_{01} = -2.62, p < .001$ ) suggests that women with higher total dynamic risk scores at initial assessment evidenced a greater decrease in scores over time relative to women with lower total dynamic risk scores at initial assessment.

*Conditional growth model: Variation as a function of Level 2 predictors.* As seen in Table 2, initial total dynamic risk scores did not vary as a function of age or Indigenous status; however, scores did vary as a function of total static risk score ( $\gamma_{03} = 0.64, SE = 0.01, p < .001$ ). Similarly, change in total dynamic risk scores did not vary as a function of age and Indigenous status; however, these scores did change over time as a function of total static risk scores ( $\gamma_{13} = -0.01, SE = 0.00, p < .001$ ); although, the magnitude of this effect was extremely small.

*Pseudo-R<sup>2</sup> results.* As described in Table 2, less than 1% of variance is explained by the unconditional growth model, whereas 36% of variance is explained by the conditional growth model. For the unconditional growth model, the proportion of within-offender variation explained by time is 45.9%, whereas for the conditional growth model, the proportion of within-individual variation explained by time is 52.3%. Adding Level 2 predictors

**TABLE 2: Two-level Linear Growth Models Examining Changes in Total Dynamic Risk Scores Over Time**

Model components	Unconditional means		Unconditional growth		Conditional growth	
	Estimate	(SE)	Estimate	(SE)	Estimate	(SE)
<b>Fixed effects initial status</b>						
Intercept ( $\gamma_{00}$ )	18.30***	(0.27)	18.54***	(0.29)	19.20***	(0.29)
Age ( $\gamma_{01}$ )	—	—	—	—	0.02	(0.02)
Indigenous ( $\gamma_{02}$ )	—	—	—	—	0.88	(0.54)
Static risk ( $\gamma_{03}$ )	—	—	—	—	0.64***	(0.01)
<b>Rate of change</b>						
Intercept ( $\gamma_{10}$ )	—	—	-0.04**	(0.01)	-0.11***	(0.01)
Time × Age ( $\gamma_{11}$ )	—	—	—	—	0.00	(0.00)
Time × Indigenous ( $\gamma_{12}$ )	—	—	—	—	0.03	(0.03)
Time × Static Risk ( $\gamma_{13}$ )	—	—	—	—	-0.01***	(0.00)
<b>Variance components</b>						
<b>Level 1</b>						
Within person ( $\sigma^2$ )	35.01***	(0.57)	18.93***	(0.38)	16.69***	(0.33)
<b>Level 2</b>						
Initial status ( $\tau_{00}$ )	201.93***	(5.59)	231.54***	(6.46)	150.76***	(4.31)
Rate of change ( $\tau_{11}$ )	—	—	0.29***	(0.01)	0.22***	(0.01)
Covariance ( $\tau_{01}$ )	—	—	-2.62***	(0.21)	-2.51***	(0.16)
<b>Fit indices</b>						
AIC	74,989.9		73,269.2		70,707.6	
BIC	75,007.8		73,304.9		70,779.1	
-2 loglikelihood	74,983.9		73,257.2		70,683.6	
<b>Pseudo-<math>R^2</math></b>						
Total outcome	—		.003		.360	
Within-person residual	—		.459		.523	
Intercept	—		—		.349	
Slope	—		—		.241	

Note. Unconditional growth model assesses within-individual effects (Level 1). Conditional growth model includes between-individual effects (Level 2 predictors). Estimate = regression coefficient; SE = standard error; AIC = Akaike information criterion; BIC = Bayesian information criterion; Estimation Method = Maximum Likelihood; Satterthwaite degrees of freedom.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

further explained outcome variation for the initial (34.9%) and change in total dynamic risk scores (24.1%).

**Dynamic Strength**

The initial assessment SPIn total dynamic strength scores ranged from 0 to 87 ( $M = 26.84$ ,  $SD = 16.51$ ). Subsequent SPIn assessments occurred at intervals spanning 3 to 30 months and the scores ranged from 0 to 89 ( $M = 27.35$ ,  $SD = 16.32$ ). The loglikelihood value associated with the random effects linear function (70,578.3) was smaller than the quadratic function (72,423.4). As such, the likelihood ratio test could not be run, indicating that the linear random effects trend best fit the data. In addition, the ICC was calculated and was found to be .912 (246.03 / [23.72 + 246.03]), indicating that 91.2% of the variance in initial total dynamic strength scores can be explained by differences among justice-involved women (see Table 3).

**TABLE 3: Two-Level Linear Growth Models Examining Changes in Total Dynamic Strength Scores Over Time**

Model components	Unconditional means		Unconditional growth		Conditional growth	
	Estimate	(SE)	Estimate	(SE)	Estimate	(SE)
<b>Fixed effects initial status</b>						
Intercept ( $\gamma_{00}$ )	27.16***	(0.30)	26.56***	(0.30)	26.29***	(0.37)
Age ( $\gamma_{01}$ )	—	—	—	—	0.06*	(0.03)
Indigenous ( $\gamma_{02}$ )	—	—	—	—	0.14	(0.69)
Static risk ( $\gamma_{03}$ )	—	—	—	—	-0.16***	(0.01)
<b>Rate of change</b>						
Intercept ( $\gamma_{10}$ )	—	—	0.08***	(0.01)	0.11***	(0.01)
Time $\times$ Age ( $\gamma_{11}$ )	—	—	—	—	-0.003*	(0.00)
Time $\times$ Indigenous ( $\gamma_{12}$ )	—	—	—	—	-0.06**	(0.02)
Time $\times$ Static Risk ( $\gamma_{13}$ )	—	—	—	—	0.002**	(0.00)
<b>Variance components</b>						
<b>Level 1</b>						
Within person ( $\sigma^2$ )	23.72***	(0.39)	12.05***	(0.24)	11.80***	(0.24)
<b>Level 2</b>						
Initial status ( $\tau_{00}$ )	246.03***	(6.67)	256.58***	(6.99)	260.01***	(7.14)
Rate of change ( $\tau_{11}$ )	—	—	0.20***	(0.01)	0.20***	(0.01)
Covariance ( $\tau_{01}$ )	—	—	-1.28***	(0.18)	-1.37***	(0.18)
<b>Fit indices</b>						
AIC	72,585.1		70,530.0		70,382.5	
BIC	72,603.0		70,565.8		70,454.0	
-2 loglikelihood	72,579.1		70,518.0		70,358.5	
<b>Pseudo-<math>R^2</math></b>						
Total outcome	—		.001		.005	
Within-person residual	—		.492		.503	
Intercept	—		—		.010	
Slope	—		—		.001	

Note. Unconditional growth model assesses within-individual effects (Level 1). Conditional growth model includes between-individual effects (Level 2 predictors). Estimate = regression coefficient; SE = standard error; AIC = Akaike information criterion; BIC = Bayesian information criterion; Estimation Method = Maximum Likelihood; Satterthwaite degrees of freedom.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

*Unconditional growth model: Level 1 change over time.* As seen in Table 3, results indicated that, on average, the initial total dynamic strength score was 26.56 and the average rate of change was an increase in scores by 0.08 points per month while on community supervision. The variance in initial status ( $\tau_{00} = 256.58$ ) and rate of change ( $\tau_{11} = 0.20$ ) were both significantly different from 0 ( $p < .001$ ), indicating that there was significant variability in both initial total dynamic strength scores and in the rate of change on these scores. The covariance between initial score and change over time ( $\tau_{01} = -1.28, p < .001$ ) suggests that women with lower total dynamic strength scores at initial assessment evidenced a greater increase in scores over time relative to women with higher dynamic strength scores at initial assessment.

*Conditional model: Variation as a function of Level 2 predictors.* As seen in Table 3, initial total dynamic strength scores did not vary as a function of Indigenous status; however, scores did vary as a function of age ( $\gamma_{01} = 0.06, SE = 0.03, p = .037$ ), whereby those who

**TABLE 4: Logistic Regression Results: Predicting Criminal Outcome Using Change in Total Dynamic Risk Scores**

Effect	<i>b</i>	<i>SE</i>	$\chi^2$	<i>p</i>	OR	95% CI
Any new charge						
Initial score	.012	0.004	8.79	.003	1.012	[1.004, 1.020]
Change over time	.289	0.133	4.70	.030	1.335	[1.028, 1.734]
New violent charge						
Initial score	.013	0.006	5.93	.015	1.014	[1.003, 1.025]
Change over time	.252	0.185	1.86	.173	1.287	[0.895, 1.850]
Technical violation						
Initial score	.027	0.004	39.48	<.001	1.027	[1.018, 1.035]
Change over time	.677	0.143	22.56	<.001	1.967	[1.488, 2.601]

Note. *b* = beta coefficient; *SE* = standard error;  $\chi^2$  = chi-square; OR = odds ratio; CI = confidence interval.

were younger had lower initial dynamic strength scores than those who were older. Initial total dynamic strength scores also varied as a function of total static risk score ( $\gamma_{03} = -0.16$ ,  $SE = 0.01$ ,  $p < .001$ ), where those who had higher total static risk scores had lower initial total dynamic strength scores relative to those who had lower total static risk scores. Change in total dynamic strength scores, however, was found to vary as a function Indigenous status ( $\gamma_{12} = -0.06$ ,  $SE = 0.02$ ,  $p = .006$ ). That is, while Indigenous and non-Indigenous women had similar initial dynamic strength scores, these scores increased at faster rates over time for non-Indigenous women, albeit the magnitude of this effect was small.

*Pseudo-R<sup>2</sup> results.* As described in Table 3, results indicated that less than 1% of variance is explained by the unconditional growth model; however, the conditional growth model did not explain any additional variance (less than 1% of variance remains explained by the conditional growth model). For the unconditional growth model, the proportion of within-offender variation explained by time is 49.2%, whereas for the conditional growth model, the proportion of within-individual variation explained by time is 50.3%. Finally, adding Level 2 predictors was not found to further explain unexplained outcome variation for both the initial total dynamic strength scores (1.3%) and the total dynamic strength scores over time (less than 1%).

#### USING CHANGE IN DYNAMIC RISKS AND STRENGTHS TO PREDICT OUTCOMES

##### Logistic Regression Results—Total Dynamic Risk Scores

The results of each of the four logistic regression analyses are presented narratively in turn and in Table 4. Results indicated that initial dynamic risk score was a significant predictor of new charges (odds ratio [OR] = 1.012, 95% CI = [1.004, 1.020]), violent charges (OR = 1.014, 95% CI = [1.003, 1.025]), and technical violations (OR = 1.027, 95% CI = [1.018, 1.035]), whereby those with higher scores were more likely to reoffend across all three outcomes. In contrast, change in dynamic risk scores was a significant predictor of new charges (OR = 1.335, 95% CI = [1.028, 1.734]) and technical violations (OR = 1.967, 95% CI = [1.488, 2.601]). This indicates that those whose scores increased were more likely to be charged for a new offense or receive a technical violation than those whose dynamic risk scores decreased over time. Change in dynamic risk scores,

**TABLE 5: Logistic Regression Results: Predicting Criminal Outcome Using Change in Total Dynamic Strength Scores**

Effect	<i>b</i>	<i>SE</i>	$\chi^2$	<i>p</i>	OR	95% CI
Any new charge						
Initial score	-.005	0.003	3.68	.055	0.995	[0.989, 1.000]
Change over time	.080	0.128	0.39	.531	1.083	[0.844, 1.391]
New violent charge						
Initial score	-.009	0.004	4.73	.030	0.991	[0.983, 0.999]
Change over time	-.289	0.193	2.24	.135	0.749	[0.513, 1.094]
Technical violation						
Initial score	-.008	0.003	6.09	.014	0.992	[0.986, 0.998]
Change over time	-.609	0.152	16.18	<.001	0.544	[0.404, 0.732]

Note. *b* = beta coefficient; *SE* = standard error;  $\chi^2$  = chi-square; OR = odds ratio; CI = confidence Interval.

however, was not a significant predictor of new violent charges (OR = 1.287, 95% CI = [0.895, 1.850]).

#### Logistic Regression Results—Total Dynamic Strength Scores

The results of the logistic regressions for each outcome are presented narratively and in Table 5. Results indicated that the initial dynamic strength score predicted new charges (OR = 0.995, 95% CI = [0.989, 1.000]), violent charges (OR = 0.991, 95% CI = [0.983, 0.999]), and technical violations (OR = 0.992, 95% CI = [0.986, 0.998]). That is, those with higher strength scores were less likely to reoffend across all outcomes. In contrast, change in dynamic strength scores did not predict new charges (OR = 1.083, 95% CI = [0.844, 1.391]) or violent charges (OR = 0.749, 95% CI = [0.513, 1.094]). However, change in dynamic strength scores predicted technical violations (OR = 0.544, 95% CI = [0.404, 0.732]), whereby those whose scores increased were less likely to reoffend.

## DISCUSSION

This study was the first to examine changes in dynamic risk factors and strength factors over time among women, and to examine how these changes relate to various criminal outcomes. Previous multiwave studies incorporating three or more timepoints to assess change in dynamic risk factors over time have focused predominately on men (e.g., Brown et al., 2009) with one exception, Greiner et al. (2015). To date, only two studies have assessed the dynamic nature of strengths and how these changes are related to offending outcomes (Hanby, 2013; Lloyd et al., 2020), although no studies have assessed how strengths change over time specifically for women. To address this gap, the two goals of this study were to: (a) examine how total dynamic risk and strength scores change over time for women, and (b) examine how these changes in dynamic risk and strength scores predict criminal outcomes.

#### PATTERNS OF CHANGE IN TOTAL DYNAMIC RISK SCORES

It was found that dynamic risk scores were related to static risk scores (i.e., those who had higher static risk scores were also higher on initial dynamic risk score). In terms of change

scores, the results indicated that total dynamic risk scores decreased over time; however, the rate of change was minimal. Those who had a higher static risk score (in addition to having a higher initial dynamic risk score) had scores on dynamic risk that decreased faster over time than those who had lower static risk scores. In fact, women who had lower static risk scores had low initial scores on dynamic risk, and these scores slightly increased over time.

There may be a number of reasons for this counter-intuitive finding. First, the majority of the women included in the analyses were low-risk cases (e.g., 57.6% scored low on static risk). Thus, given that the average static risk score for the full sample was already quite low, those who scored one standard deviation below the mean would be considered very low risk. As such, there is a possible floor effect whereby the ability of the SPIn to measure changes below this point may be difficult and any slight deviation (e.g., increase) in dynamic risk scores for this very low-risk group may result in what appears as the individual getting riskier over time.

This finding is consistent with past research indicating that while the majority of low-risk offenders do not change, when they do, it is often an increase in risk rather than a decrease. For example, Cohen et al. (2016) examined changes in risk classification levels for offenders across different risk levels. Notably, among 21,732 offenders who were initially deemed low risk, 92% were found to have no change in classification level over time. Although it was not possible for offenders in the low-risk category to decrease any further (low risk is the lowest classification they could have received), it was interesting to see that 8% of the sample increased in risk level over time. While 82% of the Cohen et al. (2016) sample were male, this floor effect may be more problematic for women given the higher prevalence of low-risk classifications. Regardless, in the current study, the overall trend indicated a decrease in dynamic risk over time for women.

Another consideration is that the exhibited change over time may be a result of raters having more information about the women at later time periods, thus having a better indication of their true score on the SPIn as time goes on. For instance, it is possible that at the start of supervision, there may have been very limited information accessible to probation officers, and thus chose to rate the various SPIn items more cautiously until additional information was collected (Lloyd et al., 2020). This would suggest that any change depicted in the findings may not reflect true change in scores, but rather indicate that assessment accuracy increases over time and that later assessments would be most predictive.

#### **PATTERNS OF CHANGE IN TOTAL DYNAMIC STRENGTH SCORES**

Results indicated that total dynamic strength scores increased over time, albeit the rate of change was minimal. Both age and static risk score significantly influenced scores on initial dynamic strength scores. Static risk total score, age, and Indigenous status each explained significant variability in changes in dynamic strength scores over time. Indigenous status explained the most variability in change in dynamic strength scores, where the rate of change in scores (increase over time) was slightly faster for non-Indigenous women. This finding could be due to missing strength factors that are important to Indigenous women but that are not captured by the SPIn, such as items related to culture, spirituality, and heritage.

Gradual change in scores for both total dynamic risk and strength could be due to the high proportion of low-risk women. It could also be that probation officers use scores from previous assessments to prepopulate scores input into the computerized SPIn software.

Personal communications with the Programs and Policy Development unit from Alberta Justice and Solicitor General (January 28, 2020) indicated that the SPIn has a “carry-forward” function often used for reassessments during ongoing supervision, allowing probation officers to focus on the risk/need/strength changes when reassessing clients. This function is useful for saving time for the probation officer but may increase the likelihood of maintaining scores across timepoints. Given that the average caseload for probation officers in Alberta is 71 clients (Bonta et al., 2019) and the large number of SPIn items, it is probable that this function is used regularly.

Nonetheless, the results were similar to those found by Greiner and colleagues (2015), which was the only study to look exclusively at change scores for women, indicating that dynamic risk scores decreased over time. The pseudo- $R^2$  results indicated that a large proportion of variance was explained by the variables included in the dynamic risk model—static risk, age, and Indigenous status. However, there may be additional variables that are pertinent to explaining the remaining variability in both the dynamic risk and strength models, such as program information or probation officer conscientiousness.

#### PREDICTION RESULTS

The second goal was to examine the predictive ability of change in total dynamic risk and strength scores. Results indicated that the initial total dynamic risk score was predictive of all three outcomes; however, the change in scores was only predictive of any new charges and technical violations. With the exception of violent charges, the change in risk scores over time was more predictive of reoffending outcomes than initial dynamic risk score. With respect to strength scores, results indicated that the initial total dynamic strength score was predictive of technical violations and violent charges. Change in total dynamic strength scores was only predictive of technical violations, albeit more predictive than the initial score.

Change scores may not be predictive of violent charges because of the low base rates of violent charges (e.g., 292 women or 10.1% were charged with a violent offense), coupled with the minimal change in scores over time. Nonetheless, the dynamic risk results were similar to those found by Greiner et al. (2015), whereby reassessment scores were more predictive of reoffending than initial assessments. However, the study by Greiner et al. (2015) did not look explicitly at the relationship between the rate of change in dynamic risk scores and reoffending outcomes, but rather examined the relationship between incremental change and reoffending and did not examine change in dynamic strength scores.

Results also indicated that the predictive ability of initial dynamic risk and strength scores and change in dynamic risk and strength scores was strongest for technical violations. Dynamic risk and strength scores may be most predictive of technical violations, in comparison with other criminal outcomes, because in some cases, the SPIn assessments are rated by the same individuals who would support a breach (i.e., a technical violation; Lloyd et al., 2020). That is, the supervising officer responsible for completing the assessment is, in some cases, responsible for initiating a breach. As change scores reflect average change rates on SPIn scores overcome, and as knowledge of an impending technical violation would only influence the more proximal SPIn assessment ratings, it is unlikely that this is the sole reason for the relationship between change scores and technical violations.

Nonetheless, previous research with women involved in the criminal justice system has demonstrated that several variables are predictive of criminal outcomes, including employment, unstable living conditions, substance use issues, peer relationships (particularly with romantic partners), and relationships with family and children (Blanchette & Brown, 2006; Van Voorhis et al., 2010). Given that the total dynamic risk and strength domains combine a number of these items, it is not surprising to see that as women remain successful in the community, their risk scores decrease and their strength scores increase, and that these changes are related to reoffending outcomes. Also, because change scores consider more proximal assessment scores and trends over time, it is not surprising that the change scores in risks and strengths tend to be more predictive than the initial assessment scores. While on supervision in the community, there are many conditions (including conditions pertaining to programming, abstaining from alcohol and drug use, and staying away from antisocial peers) that must be followed and contribute to decreases in scores on various dynamic risk items and potential increases in scores on various dynamic strength items. As such, it is reasonable that changes in scores provide a more accurate reflection of the women's current risk to reoffend than the initial assessment scores.

#### **IMPLICATIONS**

Overall, these results highlight the importance of reassessing women's risk/need factors and examining change in scores over time. In most cases, the change in scores (both dynamic risk and strength) was a stronger predictor of various criminal outcomes compared to the initial assessment score on these domains. Reassessment information can assist with identifying and tracking key areas of improvement, areas in need of further intervention and programming, and case planning—such as reassessing frequency of contact with a probation officer.

#### **Dynamic Factors**

According to Brown et al. (2009), for a factor to be considered truly dynamic, it must demonstrate significant change over time and significantly relate to the prediction of recidivism. Although this study did not examine how specific SPIn dynamic content domains changed, the results indicated that the total dynamic risk score demonstrated significant change over time and significantly predicted new charges, violent charges, and technical violations. There was also evidence that total dynamic strength scores significantly changed over time and predicted technical violations. That is, the SPIn total dynamic risk and total dynamic strength scores are in fact reflective of true dynamic factors and lend support for the use of these factors to inform case management and program planning for women. Given the predictability of strengths, perhaps the inclusion of both risks and strengths in assessment tools can enhance precision and predictive nature of reoffending outcomes; however, research is needed examining the incremental validity of strengths, and changes in strength scores over time.

#### **Use of the SPIn With Women**

The SPIn is considered a gender-informed risk assessment tool that captures both gender-neutral and gender-salient needs and strengths. Researchers have highlighted the

importance of assessing the following variables, particularly among justice-involved women: self-esteem, self-efficacy, parental stress, victimization and abuse, relationship dysfunction, mental health concerns (especially depression), and poverty and homelessness (Belknap, 2015; Van Voorhis et al., 2010). Most of these items are captured within the SPIn Full Assessment, with the exception of self-esteem and self-efficacy, which can help with case management and treatment planning. Research is encouraged to examine whether future iterations of the SPIn may benefit from including assessment of these items.

### **Strengths-Based Approach**

The strengths findings, particularly with respect to the prediction of technical violations, can have practical implications for case management practices and community programming. Offering community strengths-based correctional programming and strengths-based case management planning at the start of community supervision may assist with reductions in reoffending—especially number of technical violations. Evidence suggests that incorporating a strengths-based approach to treatment while under community supervision improves psychological well-being, likelihood of remaining and participating in treatment, and reduces recidivism for women (Messina et al., 2012). As discussed in Fedock and Covington (2018), there are five key factors that are especially important for women re-entering the community and each of these factors requires the use of strengths-based approaches to address them. These factors include finding shelter, finding stable employment and financial stability, rebuilding relationships with others, developing a sense of community belonging, and building self-esteem (O'Brien, 2001). In addition, the interaction between probation officer and client should be rooted in a strengths-based approach. Morash (2010) found that supervision practices incorporating positive, strength-building approaches, like motivational interviewing, can lead to positive changes—feelings of empowerment and increased self-esteem (Fedock & Covington, 2018) which can assist with improving the lives of women in the criminal justice system.

### **Intersectionality**

Examining the intersection of gender and other potentially discriminating factors (e.g., race, disabilities) is important to consider for future studies to assist with advancing our understanding of the relationship between gender and crime (Burgess-Proctor, 2006). Because traditional feminist criminologists have focused predominately on White, middle-class women, contemporary feminist criminologists have made strides to advance the field by recognizing that the lived experiences of minority women are very different (Burgess-Proctor, 2006). Intersectionality recognizes that there are multiple forms of oppression and inequality in addition to inequality based purely on gender. Specifically, aside from gender, women can face multidimensional inequalities on account of race, social class, sexuality, disability, socioeconomic status, and other socially constructed discriminations—such as involvement in the criminal justice system (Andersen & Collins, 2004).

*Indigeneity.* In Canada, while correctional research has been conducted on Indigenous peoples and on gender differences independently, rarely has the intersection of the two been examined (Bartels, 2012). In the Canadian criminal justice system, Indigenous peoples are

vastly overrepresented. In the current study, at the time of the initial SPIn Assessment, 28.8% of women ( $n = 829/2,877$ ) identified as Indigenous, in comparison with 6.5% of the general population in Alberta (Malakieh, 2018). This overrepresentation may be due to the unique needs of Indigenous women and the lack of resources available to address these needs; this includes increased substance use, mental health issues, residential school experiences, intergenerational trauma, welfare involvement, family history of suicide, and extreme poverty and poor living conditions on reservations (Office of the Correctional Investigator, 2015). Although several of these needs may also be relevant for non-Indigenous women, the underlying reasons behind these needs may differ, or the operationalization of these concepts may be different (Helmus et al., 2012; Wilson & Gutierrez, 2014). As such, the use of general risk assessments with Indigenous peoples in the criminal justice system requires further exploration (e.g., how well items on the SPIn capture the unique circumstances and need factors for Indigenous women).

In this study, Indigenous status was included as a covariate—given the high proportion of Indigenous women in the sample. Future research is encouraged to examine how Indigenous women change over time on both dynamic risks and strengths incorporating the examination of more culturally relevant factors. Although, for the purpose of this study, Indigenous cultures were aggregated, it is important to acknowledge the unique histories and diversity within Indigeneity; that is, there are distinct cultural, geographical, and historical differences among First Nations, Metis, and Inuit peoples that should be considered.

## LIMITATIONS

### Gender Identity Considerations

A limitation of the current study is that it focused solely on cis women (i.e., those who identify as a woman and who were labeled female at birth). In line with intersectionality, additional research is needed focusing on various gender identities and expressions, as opposed to the archaic dichotomous categorization of woman. Research and data on people who identify as gender nonconforming and transgender (GNCT) within the criminal justice system are quite limited (Woods, 2017) and the research that does exist has predominately been conducted on adolescent samples. Sexual orientation is also highly underresearched in the criminal justice field (this includes individuals who identify as lesbian, gay, bisexual, and questioning—LGBTQ+). Some reasons for the limited research in these various subgroups of those involved in the criminal justice system may, in part, be a reflection of sample size. Nonetheless, it is important to consider the unique experiences of LGBTQ+ individuals and how these experiences, both on their own and coupled with other forms of discrimination, may influence their pathways into the criminal justice system, and thus, must be captured by risk assessment and targeted in programming.

### Program Information

Program completion information was not available and, thus, the reason for any change could not be assessed. Additional change research incorporating programming information can determine whether program completion influences rates of change in dynamic risk and strength scores over time. In Alberta, programming is often completed by other areas of government or nongovernment organizations residing in the communities (personal

communications with the Programs and Policy Development unit from Alberta Justice and Solicitor General on January 28, 2020). There may be differences between various jurisdictions due to client needs and available resources. Research is needed on whether there are differences in rates of change on total dynamic risk and strength scores across various jurisdictions.

### Generalizability of Findings

To assess change over time, data on at least three timepoints were required. Because of this requirement, women who did not have three completed SPIn Full Assessments (e.g., reoffended early into supervision) were not included in analyses. However, this is a fundamental challenge of assessing change among justice-involved populations. Furthermore, as the data assessed in this study came from SPIn assessments administered by probation officers, there is a chance of desirability bias. That is, during the assessment interview, women may downplay various risk/needs to be rated more favorably by probation officers. Nonetheless, given that the SPIn assessment tool is intended for use by probation officers in real-world settings, these findings are relevant to low-risk women serving time in the community on a provincial sentence in Alberta. Additional research is needed on changes in dynamic risks and strengths overtime among higher risk samples of women.

### CONCLUSION

This study was the first to assess the predictive utility of change in dynamic risk and strength scores among justice-involved women. Results suggested that women with dynamic risk scores that decreased at faster rates over time were less likely to reoffend both in terms of technical violations and new charges. Women with dynamic strength scores that increased over time were less likely to have technical violations. Results highlight the utility of reassessment over time and provide support for the inclusion of strengths in risk assessment practices and the incorporation of a strength-based approach for women involved in the criminal justice system.

### ORCID iDs

Kayla A. Wanamaker  <https://orcid.org/0000-0002-0211-4672>

Shelley L. Brown  <https://orcid.org/0000-0002-9372-407X>

### NOTES

1. The sample from this study is subsumed within this larger sample from Brown et al. (2020).
2. See Wanamaker and Brown (2021) for a similar study with men on community supervision from Alberta, Canada.
3. In Canada, people who are sentenced for 2 years less a day, or less, enter the provincial correctional system, whereas those who are sentenced to 2 years or more enter the federal correctional system.
4. Research on the psychometric properties of the Service Planning Instrument (SPIn) is largely in its infancy. For instance, the interrater reliability of the SPIn has not been examined.
5. Importantly, the SPIn refers to strengths as protective factors, and despite preliminary findings, there has not been enough evidence to indicate that strengths assessed in the SPIn indeed demonstrate a protective effect. As such, the current study uses the term *strength* to denote positive, prosocial factors.

### REFERENCES

Andersen, M. L., & Collins, P. H. (2004). *Race, class, and gender: An anthology* (5th ed.). Wadsworth.

- Babchishin, K. M., & Hanson, R. K. (2020). Monitoring changes in risk of reoffending: A prospective study of 632 men on community supervision. *Journal of Consulting and Clinical Psychology, 88*(10), 886–898. <https://doi.org/10.1037/ccp0000601>
- Bakker, L., O'Malley, J., & Riley, D. (1998). *Storm warning: Statistical models for predicting violence*. Psychological Service, Department of Corrections.
- Bartels, L. (2012). Violent offending by and against Indigenous women. *Indigenous Law Bulletin, 8*(1), 19–22. <http://www5.austlii.edu.au/au/journals/IndigLawB/2012/27.pdf>
- Belknap, J. (2015). *The invisible woman: Gender, crime and justice* (4th ed.). Cengage Learning.
- Blanchette, K. D., & Brown, S. L. (2006). *The assessment and treatment of women offenders: An integrative perspective*. Wiley Series in Forensic Clinical Psychology.
- Bonta, J., & Andrews, D. A. (2017). *The psychology of criminal conduct* (6th ed.). Routledge Taylor & Francis.
- Bonta, J., Rugge, T., Bourgon, G., & Wanamaker, K. A. (2019). A conceptual replication of the Strategic Training Initiative in Community Supervision (STICS). *Journal of Experimental Criminology, 15*, 397–419. <https://doi.org/10.1007/s11292-019-09371-4>
- Brown, S. L., Robinson, D., Wanamaker, K. A., & Wagstaff, M. (2020). Strengths matter: Evidence from five separate cohorts of justice involved youth and adults across North America. *Criminal Justice and Behavior, 47*(11), 1428–1447. <https://doi.org/10.1177/0093854820931549>
- Brown, S. L., St, Amand, M., & Zamble, E. (2009). The dynamic prediction of criminal recidivism: A 3-wave prospective study. *Law and Human Behavior, 33*, 35–45. <https://doi.org/10.1007/s10979-008-9139-7>
- Burgess-Proctor, A. (2006). Intersections of race, class, gender, and crime: Future directions for feminist criminology. *Feminist Criminology, 1*, 27–47. <https://doi.org/10.1177/1557085105282899>
- Cohen, T. H., Lowenkamp, C. T., & Van Benschoten, S. W. (2016). Does change in risk matter? Examining whether changes in offender risk characteristics influence recidivism outcomes. *Criminology & Public Policy, 15*(2), 263–296. <https://doi.org/10.1111/1745-9133.12190>
- Dowden, C., Serin, R. C., & Blanchette, K. D. (2001). *The application of the community intervention scale to women offenders: Preliminary findings* (Research report R-97). Research Branch, Correctional Service of Canada.
- Fedock, G., & Covington, S. S. (2018). Strengths-based approaches to the treatment of female offenders. In C. Langton & J. Worling (Eds.), *The Wiley handbook of offender desistance from aggression and crime: Theory, research, and evidence-based practice*. John Wiley & Sons. [https://www.stephaniecovington.com/site/assets/files/1514/8-30-18\\_revised\\_fedock\\_covington\\_with\\_suggested\\_edits\\_8\\_30\\_2018.pdf](https://www.stephaniecovington.com/site/assets/files/1514/8-30-18_revised_fedock_covington_with_suggested_edits_8_30_2018.pdf)
- Greiner, L., Law, M. A., & Brown, S. L. (2015). Using dynamic factors to predict recidivism among women: A four-wave prospective study. *Criminal Justice and Behavior, 42*, 457–480. <https://doi.org/10.1177/0093854814553222>
- Hanby, L. J. (2013). *The relationship between dynamic risk, protective factors, and criminal recidivism in a New Zealand offender release cohort* [Doctoral dissertation]. Carleton University. <https://doi.org/10.22215/etd/2013-09969>
- Helmus, L., Babchishin, K. M., & Blais, J. (2012). Predictive accuracy of dynamic risk factors for Indigenous and non-Indigenous sex offenders: An exploration comparison using STABLE-2007. *International Journal of Offender Therapy and Comparative Criminology, 56*, 856–876. <https://doi.org/10.1177/0306624X11414693>
- Jones, N. J., Brown, S. L., Robinson, D., & Frey, D. (2015). Incorporating strengths into quantitative assessments of criminal risk for adult offenders. *Criminal Justice and Behavior, 42*, 321–338. <https://doi.org/10.1177/0093854814547041>
- Jones, N. J., & Robinson, D. (2018). Service Planning Instrument (SPIn). In J. P. Singh, D. G. Kroner, J. S. Wormith, S. L., Desmarais, & Z. Hamilton (Eds.), *Handbook of recidivism risk/needs assessment tools* (pp. 181–198). John Wiley & Sons.
- Laws, D. R., & Ward, T. (2011). *Desistance and sexual offending: Alternatives to throwing away the keys*. Guildford Press.
- Lewis, K., Olver, M. E., & Wong, S. C. (2013). The Violence Risk Scale: Predictive validity and linking changes in risk with violent recidivism in a sample of high-risk offenders with psychopathic traits. *Assessment, 20*, 150–164. <https://doi.org/10.1177/1073191112441242>
- Lloyd, C. D., Hanson, R. K., Richards, D. K., & Serin, R. C. (2020). Reassessment improves prediction of criminal recidivism: A prospective study of 3,421 individuals in New Zealand. *Psychological Assessment, 32*(6), 568–581. <http://doi.org/10.1037/pas0000813>
- Lodewijks, H. P. B., de Ruiter, C., & Doreleijers, T. A. H. (2010). The impact of protective factors in desistance from violent offending: A study in three samples of adolescent offenders. *Journal of Interpersonal Violence, 25*, 568–587. <https://doi.org/10.1177/0886260509334403>
- Malakieh, J. (2018). *Adult and youth correctional statistics in Canada, 2016/2017* (Report No. 85-002-X). Statistics Canada. <https://www150.statcan.gc.ca/n1/pub/85-002-x/2018001/article/54972-eng.htm>
- Messina, N., Calhoun, S., & Warda, U. (2012). Gender-responsive drug court treatment: A randomized controlled trial. *Criminal Justice and Behavior, 39*(12), 1539–1558. <https://doi.org/10.1177/0093854812453913>
- Morash, M. (2010). *Women on probation and parole: A feminist critique of correctional programs and services*. Northeastern University Press.

- O'Brien, P. (2001). "Just like baking a cake": Women describe the necessary ingredients for successful reentry after incarceration. *Families in Society: The Journal of Contemporary Social Services*, 82(3), 287–295. <https://doi.org/10.1606/1044-3894.200>
- Office of the Correctional Investigator. (2015). *Annual report of the Office of the Correctional Investigator, 2014–2015*. <https://www.oci-bec.gc.ca/cnt/rpt/pdf/annrpt/annrpt20142015-eng.pdf>
- Ogloff, J. R. P., & Davis, N. R. (2004). Advances in offender assessment and rehabilitation: Contributions of the risk-needs-responsivity approach. *Psychology, Crime & Law*, 10, 229–242. <https://doi.org/10.1080/0683160410001662735>
- Orbis Partners. (2003). *Service Planning Instrument (SPIn)*. <https://www.orbispartners.com/adult-risk-assessment>
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41, 71–90. <https://doi.org/10.1016/j.dr.2016.06.004>
- Scanlan, J. M., Yesberg, J. A., Fortune, C.-A., & Polaschek, D. L. L. (2020). Predicting women's recidivism using the Dynamic Risk Assessment for Offender Re-Entry: Preliminary evidence of predictive validity with community-sentenced women using a "gender-neutral" risk measure. *Criminal Justice and Behavior*, 47(3), 251–270. <https://doi.org/10.1177/0093854819896387>
- Serin, R. (2007). *The Dynamic Risk Assessment Scale for Offender Re-Entry (DRAOR)* [Unpublished scale]. Carleton University.
- Van Voorhis, P., Wright, E. M., Salisbury, E., & Bauman, A. (2010). Women's risk factors and their contributions to existing risk/needs assessment. *Criminal Justice and Behavior*, 37, 261–288. <https://doi.org/10.1177/0093854809357442>
- Wanamaker, K. A., & Brown, S. L. (2021). The relationship between patterns of change in dynamic risk and strength scores and reoffending for men on community supervision. *Criminal Justice and Behavior*. Advanced online publication. <https://doi.org/10.1177/0093854821993512>
- Wanamaker, K. A., Jones, N. J., & Brown, S. L. (2018). Strengths-based assessments for use with forensic populations: A critical review. *International Journal of Forensic Mental Health*, 17, 202–221. <https://doi.org/10.1080/14999013.2018.1451414>
- Wardrop, K., Wanamaker, K. A., & Derksen, D. (2019). Developing a risk/need assessment tool for women offenders: A gender-informed approach. *Journal of Criminological Research, Policy and Practice*, 5(4), 264–279. <https://doi.org/10.1108/JCRPP-03-2019-0017>
- Wilson, H. A., & Gutierrez, L. (2014). Does one size fit all? A meta-analysis examining the predictive ability of the Level of Service Inventory (LSI) with Indigenous offenders. *Criminal Justice and Behavior*, 41, 196–219. <https://doi.org/10.1177/0093854813500958>
- Woods, J. B. (2017). LGBT identity and crime. *California Law Review*, 105, 667–739. <https://dx.doi.org/10.15779/Z389W08Z24>
- Yang, M., Guo, B., Olver, M. E., Polaschek, D. L. L., & Wong, S. C. P. (2017). Assessing associations between changes in risk and subsequent reoffending: An introduction to relevant statistical models. *Criminal Justice and Behavior*, 44(1), 59–84. <https://doi.org/10.1177/0093854816678648>
- Zamble, E., & Quinsey, V. L. (1997). *The process of criminal recidivism*. Cambridge University.

**Kayla A. Wanamaker** received her PhD in 2020 from the psychology department at Carleton University in Ottawa, Canada. Her doctoral research assessed patterns of change in dynamic risk and strength factors and the relationship to reoffending outcomes for justice-involved men and women. She also examined typologies of men and women on community supervision that can be used to inform service delivery and rehabilitation efforts.

**Shelley L. Brown** has been an associate professor in the Department of Psychology, Carleton University, Ottawa, Canada, since 2010. Shelley studies gender differences in complex trauma, violence, strengths, and risk assessment practices among justice-involved populations.