



## A system-wide implementation and evaluation of the Strategic Training Initiative in Community Supervision (STICS)

James Bonta<sup>a,b,\*</sup>, Guy Bourgon<sup>a,c</sup>, Tanya Rugge<sup>a</sup>, Chloe I. Pedneault<sup>a,d</sup>, Seung C. Lee<sup>a</sup>

<sup>a</sup> Public Safety Canada, 340 Laurier Ave. W., Ottawa, ON K1A 0P9, Canada

<sup>b</sup> Consultant, Corrections and Criminal Behavior, 1352 Chataway Avenue, Ottawa, ON K1H 7S3, Canada

<sup>c</sup> Clinical and Consulting Psychologist, 317 Catherine Street, Ottawa, ON K1R 5T4, Canada

<sup>d</sup> Carleton University, 1125 Colonel By Drive, Ottawa, ON K1S 5B6, Canada

### ARTICLE INFO

#### Keywords:

Strategic Training Initiative in Community Supervision (STICS)  
Community supervision  
Criminogenic needs  
Intervention skills  
Recidivism

### ABSTRACT

**Purpose:** To evaluate the province-wide implementation of the Strategic Training Initiative in Community Supervision (STICS) in British Columbia, Canada.

**Methods:** Probation officers (PO) audio recorded a sample of their supervision sessions with probationers prior to and after receiving STICS training. Training involved teaching POs to apply cognitive-behavioral interventions to address the criminogenic needs of their clients. Both the general and violent recidivism rates of probationers were measured using a national criminal history database. Recidivism was defined as a new conviction.

**Results:** After training, audio recordings showed clear improvements in the content of discussions and the intervention techniques used by POs. Compared to a random sample of probationers supervised prior to STICS training, clients of STICS-trained POs had significantly lower 2-year general recidivism rates (43.0% vs. 61.4%;  $OR = 0.48$ , 95% CI [0.37, 0.61]) and violent recidivism rates (14.9% vs. 21.2%;  $OR = 0.65$ , 95% CI [0.48, 0.89]).

**Conclusions:** Training POs in how to address criminogenic needs using cognitive-behavioral interventions was associated with significant reductions in client recidivism. This project demonstrated that a community corrections agency can benefit from an evidence-based approach to supervision.

Community supervision is one of the most common criminal justice sanctions in Canada. On any given day, there are approximately 100,000 adults on community supervision, representing nearly 75% of the total correctional population (Malakieh, 2020). The United States, in 2018, had more than 4.4 million offenders under community supervision representing over 65% of the total offender population (Kaeble & Alper, 2020). Given that most offenders are supervised in the community, community supervision is integral to the effectiveness of the correctional system. Probation and parole officers meet with clients regularly to manage their risk in the community and link them to appropriate supports. Accordingly, we would expect community supervision to lower recidivism. In line with this expectation, there is evidence that supervision can reduce recidivism under certain conditions (Smith et al., 2018). However, in 2008, a quantitative review found that community supervision only produced an average reduction in recidivism of 2% and no reduction in violent recidivism (Bonta, Rugge, Scott, Bourgon, & Yessine, 2008).

An early review of the offender treatment literature led Martinson

(1974) to conclude “that there is very little reason to hope that we have found a sure way of reducing recidivism through rehabilitation” (p. 49). Although this view was not accepted by everyone, it was not until 1990 that a framework was developed outlining the circumstances under which treatment can work. The Risk-Need-Responsivity (RNR) model, formulated by Andrews, Bonta, and Hoge (1990), has become one of the most influential models of effective correctional rehabilitation (Cullen, 2013; Wormith & Zidenberg, 2018). Three principles are at the core of the model. The *risk principle* involves matching the level of service (e.g., treatment, supervision) with the offender’s level of risk. That is, intensive services should be directed to the higher risk offenders and less intensive services to those who are low risk. In fact, providing services to low risk offenders produces, on average, only a small reduction in recidivism (Andrews & Dowden, 2006) and, in some cases, increases recidivism (Bonta, Wallace-Capretta, & Rooney, 2000; Lowenkamp, Latessa, & Holsinger, 2006). The *need principle* maintains that interventions should target criminogenic needs (i.e., dynamic risk factors associated with recidivism), such as procriminal attitudes and substance

\* Corresponding author at: 1352 Chataway Avenue, Ottawa, Ontario K1H 7S3, Canada.

E-mail address: [Jim.Bonta@gmail.com](mailto:Jim.Bonta@gmail.com) (J. Bonta).

abuse, rather than targeting noncriminogenic needs, such as self-esteem and indicators of personal distress. The *responsivity principle* aims to maximize learning by delivering interventions in a way that matches the offender's abilities, motivations, and learning style, mostly using cognitive-behavioral techniques.

Correctional treatment programs that adhere to these principles produce larger reductions in recidivism than those that do not (Bonta & Andrews, 2017). Why then do we not find community supervision to have a significant impact on recidivism? To better understand these findings, Bonta et al. (2008) analysed the audio recordings of supervision sessions between probation officers (PO) and their clients from a Canadian province. They found that the sessions typically did not follow the RNR principles. Low risk clients were seen as frequently as higher risk clients, addressing criminogenic needs during supervision was haphazard, and POs rarely used cognitive-behavioral interventions.

These findings reveal two important gaps in our knowledge of the effectiveness of community supervision. First, POs in their face-to-face interactions with clients likely do not apply the most recent evidence concerning offender rehabilitation. Second, the findings highlight the need to train POs to better adhere to the RNR principles and to evaluate the effectiveness of training.

### 1. Strategic Training Initiative in Community Supervision (STICS)

STICS was developed by researchers at Public Safety Canada in the mid2000s to improve the effectiveness of community supervision by training POs to apply the principles of RNR. First, following the risk principle, STICS training was reserved for POs supervising medium and high risk clients. Second, POs were taught to distinguish criminogenic needs from noncriminogenic needs with an emphasis on addressing procriminal attitudes. Attitudes represent a major risk factor for future criminal behavior and are considered to underlie other criminogenic needs, such as substance abuse and family/marital relationships (Bonta & Andrews, 2017). That is, POs were encouraged to focus the content of their discussions with clients on criminogenic needs rather than non-criminogenic needs. Finally, in adherence to the responsivity principle, officers were taught *how* to alter procriminal attitudes using cognitive and behavioral intervention techniques. This included, for example, teaching concrete prosocial skills through modeling techniques and cognitive restructuring (i.e., replacing procriminal attitudes with prosocial attitudes). During training, officers practiced these skills and received feedback from the trainers. Additional information on the STICS model is presented in Appendix A of the Electronic Supplementary Material.

After initial training, newly developed skills were maintained and consolidated through ongoing clinical support (Bourgon, Bonta, Rugge, & Gutierrez, 2010). Clinical support activities included personal feedback, annual refresher courses (beginning approximately a year after training), and monthly meetings at the probation office. These activities served to strengthen the specific skills taught in STICS training. Overall, STICS's goal was to increase adherence to the risk (who to treat), need (what to treat), and responsivity (how to treat) principles.

To evaluate the effectiveness of STICS, a demonstration project was conducted with POs who supervised adult offenders from three Canadian provinces (Bonta et al., 2011). The POs were randomly assigned to STICS training ( $n = 33$ ) or probation-as-usual ( $n = 19$ ). The POs were asked to audio record meetings with their medium to high risk clients ( $n = 143$ ) throughout the first six months of supervision (once at intake, then three and six months later). However, the POs were free to choose their clients for recording introducing a possible selection bias. Trained research assistants coded the recordings for topics of discussion (i.e., criminogenic needs, noncriminogenic needs, and probation conditions) and intervention skills (i.e., cognitive and behavioral techniques).

STICS-trained officers spent significantly more time discussing criminogenic needs, particularly procriminal attitudes, and less time

discussing noncriminogenic needs and probation conditions with their clients. Previous research had shown that excessive attention to the probation conditions in a supervision session was associated with increased recidivism (Bonta et al., 2008). Additionally, STICS-trained officers demonstrated higher levels of intervention skills than the control POs. Clients of the STICS trained officers had significantly lower two-year recidivism rates than those of officers who did not receive STICS training (25% vs. 40%). Moreover, regardless of experimental condition, clients who had been exposed specifically to cognitive intervention techniques had significantly lower rates of recidivism after two years than those who did not (19% vs. 37%).

More recently, an attempt was made to replicate these original findings (Bonta, Rugge, Bourgon, & Wanamaker, 2019). POs were randomly assigned to STICS training ( $n = 21$ ) or probation-as-usual ( $n = 15$ ). In this study, medium to high risk clients were randomly assigned to POs to reduce the potential for selection bias (a source of bias in the 2011 study). Following a similar methodology as in the 2011 experiment, POs audio recorded sessions with their clients. Compared to the probation-as-usual group, STICS-trained POs focused more on procriminal attitudes during their sessions and less time discussing probation conditions. On the other hand, POs in the control group spent more time discussing the criminogenic needs of family/marital and substance abuse. Additionally, STICS-trained POs showed significantly higher levels of relationship building and cognitive intervention skills when compared to the control POs.

Overall, there was no difference in two-year recidivism rates between clients supervised by the STICS-trained POs and the control POs (52% vs. 49%). However, regardless of experimental condition, the use of cognitive intervention techniques was associated with reduced recidivism. After controlling for age and risk, the average recidivism rate for clients exposed to cognitive interventions was significantly lower (43%) than for those who were not (54%). This finding is consistent with other research on community supervision where cognitive intervention techniques were used (Labrecque, Smith, Schweitzer, & Thompson, 2013; Robinson et al., 2012).

#### 1.1. Present study

The community corrections division of the province of British Columbia (BC), one of the sites of the 2011 STICS evaluation, was encouraged by the findings and decided to implement the STICS model across its jurisdiction. Therefore, whereas previous evaluations of STICS were based on small demonstration projects with volunteer POs, the current study evaluates a large scale implementation where POs were required to attend training and to use STICS skills in their supervision of probationers (i.e., most of the POs were non-volunteers). The system-wide roll-out of STICS in BC began in September 2011 with the last training occurring in February 2015. The training was provided by the developers of STICS.

The study posed two research questions. First, is STICS training associated with a change in officer behavior, particularly on discussions of criminogenic needs and the application of intervention techniques? Specifically, we expected that STICS-trained POs would spend more time discussing criminogenic needs (i.e., following the need principle) and demonstrate improved cognitive-behavioral intervention skills (i.e., responsivity principle). As all POs were to receive STICS training, we assessed PO behavior pre-and post-training (i.e., single group pre-and post-test design; Campbell & Stanley, 1966).

Second, if there are changes in PO behavior consistent with the need and responsivity principles, is there an associated decrease in client recidivism? In STICS training, a large emphasis is placed on learning how to apply cognitive-behavior interventions. However, as suggested by previous research (Bonta et al., 2011; Bonta, Rugge, Bourgon, & Wanamaker, 2019), we also hypothesized that the use of cognitive interventions, distinct from behavioral techniques, would be more tied to reduced recidivism. We evaluated the impact of STICS training and PO

behavior on client recidivism using a quasi-experimental design (Campbell & Stanley, 1966).

**Research Question 1.** Is STICS Training Associated with Changes in PO Behavior?

**2. Method**

**2.1. PO participants**

In total, 357 POs participated in the study. All POs were at the “PO24” level, meaning that they were tasked with supervising medium to high risk probationers. Probationers’ level of risk was determined using the province’s risk/need assessment instrument. Demographic information was available only for those POs who voluntarily agreed to answer items from a self-report questionnaire administered during training. Approximately 20% of POs did not complete the questionnaire. Demographic information and other PO characteristics are presented in Table 1. POs were, on average, 41 years old (range = 23–63), female (64%), and Caucasian (79%). Almost all had a college or university degree and had an average of 10 years of work experience. Forty-seven POs were coaches (one coach per office; for a very small office of only a few POs, a coach was assigned to two or three offices).

**2.2. Measures**

**2.2.1. Pre-training (baseline) audio recordings**

Prior to attending training, each PO was asked to audio record two probationers early in supervision (between one and three months) and two probationers later in supervision (three to six months). Until the end of May 2013, the selection of these clients was left to the PO (i.e., clients were not randomly assigned). Starting in June 2013, the sampling strategy was revised and clients were randomly assigned to POs by the duty officer as clients were sentenced and assigned to a probation office (a random court process). The duty officer also flagged one alternative or back up client for each officer. The alternative client was used only if the first client was unable to participate (e.g., absconds before the baseline recording or refused to be involved in the study). Unfortunately, the rate of refusals and characteristics of the clients who refused were not collected. This selection process continued until the PO had recorded one baseline session with four clients.

**Table 1**  
PO demographics and characteristics.

	%	n	/N
Gender			
Male	36.1	104	/288
Female	63.9	184	/288
Ethno-cultural background			
Caucasian	79.1	227	/287
Asian	4.2	12	/287
Aboriginal/Métis	4.2	12	/287
East Indian	8.7	25	/287
Other <sup>a</sup>	3.7	11	/287
Education			
High school	0.4	1	/285
College/Undergraduate degree	91.6	261	/285
MA Degree	8.1	23	/285
STICS coach status			
No	86.8	310	/357
Yes	13.2	47	/357
	<i>M (SD)</i>	<i>n</i>	
Age	40.93 (9.27)	283	
Years of experience	10.02 (7.08)	287	

Notes. *M* = mean. *SD* = standard deviation. *Ns* vary as not all information was available.

<sup>a</sup> Other included one person each in the following categories: Black, Hispanic, Caucasian/Aboriginal, Caucasian/Asian, East Indian/Caucasian, Middle East, Sikh, South Asian, South East Asian, “mixed”, and not specified.

**2.2.2. Post-training audio recordings**

As new clients were admitted to the office after sentencing, the duty officer assigned the client to a trained PO. Each officer received one STICS client per month (one alternative STICS client per month was also identified). This selection process continued each month until the PO had recorded at least one session with six clients (two medium risk and four high-risk clients). The PO conducted the first post-training recording with a client (i.e., Tape 1) within 90 days from the start date of the client’s supervision order. Tape 2 occurred three months after the first tape, and Tape 3 was completed three months after Tape 2.

**2.3. Selecting recordings for analysis**

Over 2600 audio recordings were submitted by POs (913 pre-training/baseline recordings and 1756 post-training recordings). Resources did not permit coding all the recordings. Therefore, to ensure a representative sample of PO behavior, at least one recording was randomly selected from each client. In total, 1236 recordings from 1112 individual clients were coded (283 baseline and 953 post-training recordings). This selection process resulted in a median of two audio recordings per PO (*M* = 2.60, range = 1–9). Seven hundred and twenty-one clients had one recording, 92 clients had two recordings, and 16 had three.

**2.4. Coding methodology**

The audio recordings were coded according to the content of the discussions and the skills demonstrated by the POs. The coding methodology was identical to that used by Bonta et al. (2011); Bonta, Ruggie, Bourgon, & Wanamaker (2019). Teams of two coders listened to each five-minute segment of the recording and coded each segment for the presence or absence of specific discussion content areas and skills (described next; see Appendix A of the Electronic Supplementary Material for a more detailed description). Each coder conducted the scoring independently and then reached a consensus final score, which was used for analyses. Scoring disagreements were discussed with a senior researcher for a final decision.

Fifty-eight audio recordings (10% of available audio recordings available at the time) were randomly selected for interrater reliability analysis. The Intraclass Correlation Coefficient (ICC; one-way random effect model) was calculated for continuous (i.e., ordinal) items and Cohen’s Kappa for categorical (e.g., yes/no) items. Cicchetti (1994; p. 286) has stated that when the ICC is between 0.60 and 0.74, “the level of clinical significance is good”. Therefore, an ICC value of 0.60 was set as the minimum criterion for including ordinal items in further analyses. Kappa was set at a value of 0.81 or higher (“almost perfect”; Landis & Koch, 1977, p. 165) as the minimum criterion for categorical items.

**2.4.1. Discussion content areas**

First, all audio recordings were coded for the presence or absence of discussions surrounding criminogenic needs as well as probation conditions and noncriminogenic needs. The criminogenic needs that were assessed and met the minimum ICC criterion were the following: pro-criminal attitudes, procriminal peers, family/marital, employment/education, and substance abuse. The ICC results for all the discussion areas can be found in the Electronic Supplementary Material (Table C1 of Appendix A). The average ICC for the discussion content areas was 0.77 with all areas meeting the minimum ICC criterion except for antisocial personality pattern (ICC = 0.43); therefore, we excluded this content area from analyses. Because the provincial risk/need assessment instrument did not measure the client’s involvement in prosocial leisure and recreational activities, the criminogenic need of leisure/recreation was not coded.

**2.4.2. Intervention skills**

Intervention skills were rated as absent or present (all Kappas were

0.84 or higher). Next, if the skill was present, the quality of the skill was rated on a 7-point scale from *low-quality* (1) to *high-quality* (7). These quality rating scores were used to assess changes in PO behavior skills with each intervention skill ranging from 0 (absent) to 7 (high-quality). The first intervention skill that was constructed requires comment. In the original STICS evaluation, it was referred to as “relationship skills,” but three out of the five “relationship skills” items did not meet minimum ICC criterion (Bonta et al., 2011 did not employ a minimum ICC criterion for the inclusion of items). The two items that met the criterion were “role clarification” and “agreement on goals”, which were labelled “collaborative skills” in the current study.

The second intervention skill was session structuring and reflects the steps that should be followed in almost all supervision sessions according to the STICS model. Next, we examined cognitive and behavioral intervention skills. Cognitive interventions refer to those that target thinking and use cognitive restructuring techniques to replace procriminal thinking with prosocial thinking. Behavioral intervention techniques included, for example, the use of effective reinforcement and practicing new behaviors either within session or as a homework assignment.

A composite score for cognitive-behavioral techniques was created by summing total scores for cognitive and behavioral skills. Additionally, a second composite category labelled “effective correctional skills” encompassed all intervention skills described above. Interrater reliabilities for the skills are presented in the Electronic Supplementary Material (Appendix A, Table C2).

### 2.5. Analytic plan

Our first research question examined whether STICS training was associated with changes in PO behavior, including differences in the content of discussions with clients and the quality of intervention skills. To address this question, POs’ pre-and post-training audio recordings were compared to determine within-person changes in the proportion of time spent discussing criminogenic and noncriminogenic needs, as well as the quality of intervention skills. Therefore, only POs who submitted both pre-and post-training audio recordings were included in these analyses.

The proportion of 5-min segments during which a content area was discussed was computed by dividing the number of 5-min segments the topic was discussed by the total number of 5-min segments within a session. The proportion of each session a topic was discussed was then aggregated at the PO level for pre-and post-training audio recordings separately. For example, if a PO submitted more than one post-training audio recording, we averaged proportions for each content area over all post-training recordings, resulting in a single mean proportion for each content area. Similarly, intervention skills were aggregated at the PO level for pre-and post-training audio recordings separately.

For all comparisons, Cohen’s *d* was computed as a measure of effect size as follows:

$$d = (M_1 - M_2) / \sqrt{\frac{(n_1 - 1)SD_1^2 + (n_2 - 1)SD_2^2}{n_1 + n_2 - 2}}$$

where  $M_1$  represents the mean for the post-training recording,  $M_2$  represents the mean for the pre-training recording,  $n_1$  and  $n_2$  are the sample sizes for post- and pre-training recordings, and  $SD_1$  and  $SD_2$  are the standard deviations for post- and pre-training recordings, respectively. Positive *d* values indicate an increase in discussion content or improvement in intervention skills. Cohen (1992) proposed heuristic values for interpreting *d*, with 0.2 corresponding to a small effect, 0.5 to a medium effect, and 0.8 to a large effect. Also computed was the 95% confidence interval (CI) for *d*. Confidence intervals that include zero indicate no statistically significant difference.

## 3. Results

Of the 357 POs who participated in this study, 201 submitted samples of both pre-training and post-training audio recordings. On average, the post-training recordings ( $M = 24.55$ ,  $SD = 8.60$ ) were significantly longer than the pre-training recordings ( $M = 18.60$ ,  $SD = 9.52$ ,  $d = 0.66$ , 95% CI [0.46, 0.86]). As noted earlier, an individual PO could record more than one client at baseline and after training. Thus, a total of 730 clients were supervised by these 201 POs. There were no significant differences in the personal demographics of this sub-group of probationers and the overall sample of probationers described in Table 2. Although Table 2 shows some missing data, the rate was less than 7%. Thus, the potential for bias was minimal and no specialized methods were used to account for missing data.

The mean proportion of 5-min segments during which a topic was discussed pre-and post-training is presented in Table 3. On average, POs discussed noncriminogenic needs and probation conditions significantly less frequently after STICS training. Because attitudes were the primary focus of STICS training, we examined the extent to which POs addressed attitudes in each session, regardless of whether they identified attitudes as a criminogenic need for the client. A large difference was observed from pre- to post-training for attitudes, with POs discussing attitudes with their clients for a significantly higher proportion of each session after STICS training.

For the remaining criminogenic needs, only clients for whom the criminogenic need was identified by the risk/need assessment were examined. This reduced the sample size for each criminogenic need because only POs who identified the criminogenic need for at least one client pre-and post-training were included in these analyses. A higher proportion of each session was focused on identified problem attitudes at post-training compared to pre-training (Table 3). Small to moderate effect sizes were observed for family/marital, employment/education, and substance abuse, such that these criminogenic needs were discussed less frequently post-training; however, the 95% confidence intervals for family/marital and employment/education included zero indicating a nonsignificant difference. Discussions of procriminal peers did not differ regardless of training.

Similar to discussion areas, within-PO changes in intervention skills from pre- to post-training were also assessed. Mean pre-training and post-training scores on each skill are shown in Table 4. Moderate to large effect sizes were observed across all intervention skills. Specifically, the quality of collaborative skills, session structuring skills, and cognitive and behavioral intervention techniques significantly improved after training. Notably, the average quality rating for total effective correctional skills almost doubled from pre- to post-training.

**Research Question 2.** Is STICS Intervention Associated with Reduced Client Recidivism?

## 4. Method

### 4.1. Client samples

At first glance, it would appear that an obvious group to compare the recidivism rates of the 829 post-training clients (i.e., R-STICS group) would be the 283 clients who were recorded before the PO attended STICS training (Recorded [R]-Baseline group). However, there were two problems with this comparison. First, as a result of the short timeframe within which all POs had to receive STICS training, 65.7% of R-Baseline clients may have been exposed to STICS to some extent because their probation order extended beyond the date of their PO’s STICS training, and they did not recidivate prior to the PO’s training. Second, there was likely a selection bias stemming from two sources: R-Baseline clients in the early implementation stages were not randomly selected, and clients who agreed to be recorded may not well represent the probation population (i.e., a self-selection factor). Indeed, analysis of risk scores (to be

**Table 2**  
Demographics and characteristics of R-STICS (n = 829), R-Baseline (n = 283), and NR-Baseline (n = 467) Clients.

	R-STICS			R-Baseline			NR-Baseline			R-STICS vs. R-Baseline	R-STICS vs. NR-Baseline
	%	n	/N	%	n	/N	%	n	/N	Cramer's V	
Gender										.04 <sup>ns</sup>	.05 <sup>ns</sup>
Male	84.4	693	/821	81.3	213	/262	87.7	407	/464		
Female	15.6	128	/821	18.7	49	/262	12.3	57	/464		
Ethno-cultural background										.05 <sup>ns</sup>	.04 <sup>ns</sup>
Caucasian	62.3	512	/822	65.9	182	/276	59.2	273	/461		
Indigenous	28.6	235	/822	23.9	66	/276	29.5	136	/461		
Other <sup>a</sup>	9.1	75	/822	10.1	28	/276	11.3	52	/461		
Marital status										.06 <sup>ns</sup>	.05 <sup>ns</sup>
Single	66.3	536	/809	71.7	195	/272	68.7	311	/453		
Married or common law	22.9	185	/809	17.3	47	/272	18.5	84	/453		
Separated/divorced/widowed	10.9	88	/809	11.0	30	/272	12.8	58	/453		
Education										.03 <sup>ns</sup>	.05 <sup>ns</sup>
No high school degree	53.4	425	/796	49.6	132	/266	58.5	261	/446		
High school degree	34.4	274	/796	36.5	97	/266	30.7	137	/446		
College/university	12.2	97	/796	13.9	37	/266	10.8	48	/446		
Index offence type										.04 <sup>ns</sup>	.06 <sup>ns</sup>
Violent	47.6	410	/827	45.2	128	/283	50.9	222	/436		
Non-violent	38.5	318	/827	41.0	116	/283	40.8	178	/436		
Administration of justice	12.0	99	/827	13.8	39	/283	8.3	36	/436		
Any history of sexual crimes										.02 <sup>ns</sup>	.01 <sup>ns</sup>
Yes	6.5	54	/829	5.7	16	/283	5.8	24	/464		
No	93.5	775	/829	94.3	267	/283	94.2	437	/464		
	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>		Cohen's <i>d</i> [95% CI]	
Age (year)	36.59	10.55		35.05	11.52		35.34	10.21		0.14 [0.01, 0.28]	0.12 [0.01, 0.23]
LSI-R (Criminal History sub-scale)	5.56	2.27		4.62	2.62		5.06	2.53		0.41 [0.26, 0.53]	0.21 [0.10, 0.32]
Follow-up time (year)	3.83	1.04		4.21	0.83		2.97	1.03		0.39 [0.25, 0.53]	0.82 [0.71, 0.94]

Notes. R-STICS = Recorded STICS clients. R-Baseline = Recorded baseline clients. NR-Baseline = Non-recorded baseline clients. LSI-R = Level of Service Inventory-Revised. Information was based on the electronic probation records. Cramer's *v* is a measure of the strength of association between two categorical variables (0 ≤ *v* ≤ 1); *v* < 0.10 indicates that the strength of association from chi-square analyses is negligible.

<sup>a</sup> Other ethno-cultural groups included East Indian, Asian, Black, and Hispanic individuals.

**Table 3**  
Within-PO changes in discussion areas from pre- to post-training (proportion of session).

Discussion area	<i>n</i>	Pre-training <i>M</i> ( <i>SD</i> )	Post-training <i>M</i> ( <i>SD</i> )	Cohen's <i>d</i> [95% CI]
Probation conditions	201	0.38 (0.32)	0.30 (0.22)	<b>-0.29</b> [-0.49, -0.09]
Noncriminogenic	201	0.51 (0.29)	0.36 (0.21)	<b>-0.59</b> [-0.79, -0.39]
Probation conditions and noncriminogenic needs	201	0.60 (0.26)	0.51 (0.21)	<b>-0.38</b> [-0.58, -0.18]
Any and all criminogenic needs	201	0.39 (0.25)	0.47 (0.21)	<b>0.35 [0.15, 0.54]</b>
Attitudes – All sessions	201	0.02 (0.10)	0.18 (0.18)	<b>1.10 [0.89, 1.31]</b>
Criminogenic need identified for the client <sup>a</sup>				
Attitudess	51	0.02 (0.10)	0.18 (0.20)	<b>1.01 [0.60, 1.42]</b>
Peers	123	0.18 (0.22)	0.17 (0.18)	-0.05 [-0.30, 0.20]
Family/marital	84	0.27 (0.25)	0.21 (0.20)	-0.27 [-0.57, 0.04]
Employment/education	74	0.28 (0.27)	0.21 (0.19)	-0.30 [-0.62, 0.02]
Substance abuse	148	0.35 (0.31)	0.22 (0.21)	<b>-0.49</b> [-0.72, -0.26]

Notes. *M* = mean. *SD* = standard deviation. Values are proportions (range is 0 to 1). CI = confidence interval. Bolded values indicate effect sizes with CIs that do not include zero, indicating statistical significance (*p* < .05).

<sup>a</sup> POs had to have identified the criminogenic need in at least one baseline and one post-training client to be included in these analyses.

**Table 4**  
Within-PO changes in intervention skill scores from pre- to post-training.

Skill	Range	<i>n</i>	Pre-training <i>M</i> ( <i>SD</i> )	Post-training <i>M</i> ( <i>SD</i> )	Cohen's <i>d</i> [95% CI]
Collaborative	0–14	201	0.72 (1.43)	2.27 (1.99)	<b>0.89 [0.69, 1.10]</b>
Structuring	0–35	200	6.77 (3.22)	9.55 (4.09)	<b>0.76 [0.55, 0.96]</b>
Cognitive	0–28	201	0.03 (0.16)	1.71 (2.36)	<b>1.00 [0.80, 1.21]</b>
Behavioral	0–35	200	3.54 (2.41)	6.62 (3.31)	<b>1.06 [0.85, 1.27]</b>
Cognitive-behavioral <sup>a</sup>	0–63	200	3.57 (2.42)	8.35 (4.82)	<b>1.25 [1.04, 1.47]</b>
Total effective correctional <sup>b</sup>	0–98	200	11.03 (5.46)	20.15 (8.65)	<b>1.26 [1.05, 1.48]</b>

Notes. *M* = mean. *SD* = standard deviation. CI = confidence interval. Bolded values indicate effect sizes with confidence intervals that do not include zero indicating statistical significance (*p* < .05).

<sup>a</sup> Sum of cognitive and behavioral skills.

<sup>b</sup> Sum of collaborative, structuring, and cognitive behavioral skills.

described later) found R-Baseline clients were lower risk than R-STICS clients (Cohen's *d* = 0.41; Table 2).

To address the issues raised in using the R-Baseline as a comparison group, we randomly selected a sample of non-recorded pre-STICS clients (NR-Baseline; *n* = 467). NR-Baseline clients were those who started supervision with a PO between 60 and 730 days before the PO attended training and whose supervision sessions were not audio recorded. Additional criteria for selecting this group were as follows: (a) PO must have been listed as a supervisor for the client at least once during this timeframe, (b) the client's supervision order had to be at least one year in duration, (c) the client must have had a risk supervision assessment

rating of medium or high, and (d) the client was not part of the R-Baseline group. The NR-Baseline clients, unlike R-Baseline clients, were not “cherry-picked” by POs for inclusion. Additionally, there was less potential exposure to STICS for NR-Baseline (40.3%) than R-Baseline clients (65.7%). Furthermore, because NR-Baseline clients were randomly selected from historical records, they were more representative of the pre-STICS population of probationers. The NR-Baseline clients were comparable to R-STICS clients in terms of personal and criminal history demographics, except for age and risk scores (see Table 2).

## 4.2. Measures

### 4.2.1. Recidivism

Our primary source of recidivism information was the national criminal history records held by the Royal Canadian Mounted Police (RCMP). The advantage of RCMP records is that they capture convictions across Canada and are not limited to a particular province. Another advantage is that the recidivism results can be compared to earlier evaluations of STICS conducted in Canada that used the same indicator of recidivism. However, RCMP records were unavailable for 74 cases. For the missing cases, a search of the provincial database enabled 72 of the 74 cases to be coded. Therefore, there were only two probationers for whom we had no follow-up information and they were treated as missing in the recidivism analysis.

Recidivism was defined as any new reconviction from the date of the client’s first audio recording for the R-STICS group, and the effective community supervision start date for the NR-Baseline group, to the end of the follow-up period (i.e., the date of receiving national and/or provincial criminal records). There were 34 clients deceased before the end of follow-up time, and therefore deceased dates were used when computing the follow-up time for these clients. As the start of the follow-up period differed across groups, we controlled for the time offence free in the community before selection into the study (i.e., the start of the follow-up period) when estimating recidivism rates.

Recidivism was separated into *general recidivism* (i.e., any new conviction) and *violent recidivism* (i.e., only violent convictions). This is the first evaluation of STICS to examine violent recidivism. Violence is not only a major concern for the public and correctional agencies, but it also answers a methodological question regarding our ability to measure the impact of STICS training on low base rate behavior. In addition, from an RNR perspective there is empirical evidence suggesting that following the principles is associated with reduced violent behavior (Dowden & Andrews, 2000).

### 4.3. Analytic plan

The second primary research question examined was whether supervision by a trained PO was associated with a reduction in client recidivism. We addressed this question in two ways. First, we compared the recidivism rates of R-STICS clients (i.e., audio recorded clients of a STICS trained PO) with the recidivism rates of NR-Baseline clients (i.e., non-audio recorded clients randomly selected from the POs caseload prior to STICS training). As previously noted, the NR-Baseline clients were more representative of the PO caseload prior to STICS training than R-Baseline clients, and they were less likely to have been exposed to a STICS-trained PO during the study period. Second, for all clients who were audio-recorded (R-STICS and R-Baseline), we examined the relationship between PO behavior during the audio-recorded sessions (i.e., discussion content and intervention skills) and client recidivism. This allowed us to explore whether some PO behaviors were more strongly associated with client recidivism than others.

Both a two-year fixed follow-up as well as varying follow-up periods were used to evaluate recidivism outcomes. For the fixed two-year follow-up, groups were compared on observed recidivism rates and also with logistic regression analysis to examine the association between

STICS training and recidivism while controlling for covariates (described later). Odds ratios ( $\exp^{B1}$ ) indicate the change in relative risk associated with a one-unit change in the predictor variables. An odds ratio below 1 indicates that R-STICS clients have lower recidivism rates. No association is indicated when the 95% CI of the odds ratio contains 1. The  $B0$  (an intercept value) represents the expected recidivism rate for the NR- or R-Baseline group in logit units, which could be converted to a probability ( $1/[1 + \exp^{(-B0)}]$ ). Also calculated were the expected recidivism rate at any specific level of a predictor variable (e.g.,  $1/[1 + \exp^{(-B0+B1)}]$ ).

Cox regression survival analysis was used to examine the effect of STICS training on recidivism rates at varying times. Hazard ratios [HR] indicate how much the rate of recidivism increases or decreases based on the predictor variables after controlling for varying follow-up times. An HR below 1 indicates that R-STICS clients (or clients with more exposure to specific content/skills) had a lower recidivism rate than NR-Baseline clients (or clients with less exposure to specific content/skills). When computing time to recidivism for survival analyses, the new offence sentencing date was used to determine the time to recidivism.

Covariates known to be associated with recidivism that differed across groups were applied in regression and survival analyses. First, the Criminal History subcomponent of the Level of Service Inventory-Revised (LSI-R; Andrews & Bonta, 1995) was used to control for the differences in offender risk between the groups. The Criminal History sub-component consists of 10 historical items scored 0 for absent and 1 for present and was based on file review. The Criminal History score predicted recidivism in our sample moderately well (AUC = 0.66; 95% CI [0.63, 0.69]). Second, we controlled for time offence-free in the community before selection into the study. This is an important confound because, as time offence-free in the community increases, the risk of recidivism decreases (regardless of the client’s initial level of risk; Flores, Holsinger, Lowenkamp, & Cohen, 2017; Hanson, Harris, Letourneau, Helmus, & Thornton, 2018). The third control was for client age at the start of the probation period (Hanson, 2002; Hirschi & Gottfredson, 1983).

To summarize, our major measure of effect size was Cohen’s  $d$  and we report 95% confidence intervals. The disadvantage of relying on  $p$  values is that it says nothing about the size of the effect and null hypothesis significance testing has a host of problems, including missing a real effect (Cumming, 2014; Gendreau & Smith, 2007). Logistic regression analysis was conducted in this study to predict categorical outcomes (i.e., reoffending or not) based on predictor variables that are continuous or categorical. Additionally, Cox regression was used to investigate the association between the survival time of clients (i.e., time without recidivism) and one or more predictor variables (e.g., intervention skills).

## 5. Results

The average length of the follow-up period was 3.5 years ( $SD = 1.1$ ). Criminal history records for the NR-Baseline clients were requested earlier from the RCMP and thus had an earlier end date. As a result, the average follow-up period for NR-Baseline clients was shorter ( $M = 2.97$  years,  $SD = 1.03$ ) than for R-STICS clients ( $M = 3.83$ ,  $SD = 1.04$ ;  $t [968.42] = 14.29$ ,  $p < .001$ ; Table 2). Overall, approximately half (52.7%) of the clients from the two groups were reconvicted of a new offence at some point within the follow-up period. The observed general recidivism rates with a 2-year fixed follow-up time were significantly lower for the R-STICS clients (43.0%) compared to the NR-Baseline clients (61.4%) as was the violent recidivism rate (14.9% vs. 21.2%; Table 5). Similar trends were observed after controlling for age, risk, and time offence-free in the community (adjusted recidivism rates; Table 5). In Table 5, we used a 2-year fixed follow-up with logistics regression when estimating the base rates ( $B0$  parameter). This allows for comparison to the results of previous studies as well as yielding odd ratios ( $B1$ ).

**Table 5**  
Observed and adjusted recidivism rates with a 2-year fixed follow-up period.

Observed Recidivism Rates			Adjusted Recidivism Rates <sup>b</sup>		
NR-Baseline % (n/N)	R-STICS % (n/N)	OR [95% CI] <sup>a</sup>	NR-Baseline % (n/N)	R-STICS % (n/N)	OR [95% CI]
<b>General</b>					
Recidivism					
61.4 (243/ 396)	43.0 (343/ 798)	<b>0.48</b> [0.37, 0.61]	60.4 (239/ 396)	42.7 (340/ 795)	<b>0.49</b> [0.37, 0.65]
<b>Violent</b>					
Recidivism					
21.2 (84/ 396)	14.9 (119/ 798)	<b>0.65</b> [0.48, 0.89]	18.9 (75/ 396)	13.5 (107/ 795)	<b>0.67</b> [0.47, 0.95]

Note. OR = odds ratio; CI = confidence interval. Numbers in bold indicate statistical significance ( $p < .05$ ).

<sup>a</sup> Odds ratios and the 95% confidence intervals were calculated using a two-by-two frequency table.

<sup>b</sup> The recidivism rates of NR-Baseline and R-STICS groups were estimated by logistic regression analysis after controlling for age (36 years old), criminal history (LSI-R; an average score of 5), and time offence-free (0 month). Logistic regression analysis was conducted based on clients with data available for all model variables.

Limitations of using fixed follow-up periods are lower power due to reduced sample size and loss of information gained from longer follow-ups. To address these limitations, we used the full follow-up time available. Specifically, we examined the relationship between group (NR-Baseline vs. R-STICS) and general and violent recidivism using Cox regression controlling for important client characteristics (see Table 6). As expected, higher risk scores on the LSI-R Criminal History subscale and younger age predicted higher general and violent recidivism rates across all clients. Additionally, more time offence-free in the community before being selected into the study predicted lower general recidivism rates (but not violent recidivism). After controlling for these variables and varying follow-up time, R-STICS clients had significantly lower general recidivism rates (HR = 0.69 [0.58, 0.83]). At any point during the follow-up period, R-STICS clients were 31% less likely to be reconvicted of any new offence compared to NR-Baseline clients (this rate was constant over time, Wald  $\chi^2$  [1] = 0.33,  $p = .86$ ). Although violent recidivism tended to be lower among R-STICS clients than NR-Baseline clients, this difference was not statistically significant (HR = 0.81 [0.61, 1.11]). Fig. 1 visually displays the survival curves for (a) general and (b) violent recidivism events for the two groups.

Next, the audio recordings from the R-Baseline and R-STICS clients allowed us to examine the association between specific PO behaviors and client recidivism. PO behavior was aggregated at the client level rather than the PO level. That is, all of a client's recordings were

**Table 6**  
Cox regression predicting general recidivism between NR-baseline and R-STICS groups.

Variables	General Recidivism		Violent Recidivism	
	HR	95% CI	HR	95% CI
NR-Baseline (N = 467) vs. R-STICS (N = 823)	<b>0.69</b>	[0.58, 0.83]	0.81	[0.61, 1.11]
Criminal history subscale (LSI-R)	<b>1.17</b>	[1.13, 1.22]	<b>1.13</b>	[1.07, 1.19]
Age (year)	<b>0.97</b>	[0.97, 0.98]	<b>0.96</b>	[0.95, 0.98]
Time offence-free <sup>a</sup> (month)	<b>0.96</b>	[0.93, 0.99]	1.00	[0.96, 1.04]

Note. HR = hazard ratio. CI = confidence interval. Numbers in bold indicate statistical significance ( $p < .05$ ).

<sup>a</sup> Month(s) between the supervision start date and the first audio recording date for R-STICS clients.

aggregated to create mean scores for exposure to certain PO behaviors, including the proportion of each session spent discussing criminogenic needs and intervention skills.

Adding proportion of time spent discussing each criminogenic need to a hierarchical Cox regression model including group (R-Baseline vs. R-STICS), age, criminal history, and time offence free significantly improved prediction of general recidivism ( $\Delta\chi^2 = 15.40$ ,  $p = .009$ ; see Table 7). Specifically, procriminal attitudes made a significant contribution to the prediction of general recidivism, such that every 10% increase in the proportion of each session spent discussing procriminal attitudes resulted in approximately a 5% decrease in recidivism. In contrast, a higher proportion of each session devoted to substance abuse was associated with significantly higher recidivism rates, such that every 10% increase in the proportion of each session spent discussing substance abuse resulted in approximately a 4% increase in recidivism. None of the other discussion areas were independently associated with general or violent recidivism.

Adding quality ratings for all intervention skills to the regression model including group (R-Baseline vs. R-STICS), age, criminal history, and time offence free significantly improved the prediction of general recidivism ( $\Delta\chi^2$  [4, N = 1102] = 11.11,  $p = .025$ ; results not in table). However, after examining each intervention skill individually, only cognitive intervention skills made a significant contribution to the model (see Table 8, Model 1). Adding the other intervention skills (collaboration, structuring, and behavioral intervention skills) to a model already including cognitive intervention skills did not significantly improve the prediction of general recidivism (see Table 8, Model 2). Additionally, their shared variance with cognitive intervention skills rendered the latter nonsignificant in Model 2 (Table 8). None of the intervention skills were significantly associated with violent recidivism.

Even when mere exposure to cognitive techniques was examined, regardless of quality and client group, the fixed two-year general recidivism rate for clients exposed to cognitive techniques (27.6%) was significantly lower than the general recidivism rate for those not exposed to cognitive techniques (35.8%; odds ratio = 0.68 [0.51, 0.92]). When the entire follow-up period for each client was used, clients who were exposed to cognitive techniques were approximately 28% less likely to be reconvicted of any new offence compared to clients who had not been exposed to cognitive techniques at all (hazard ratio = 0.72 [0.58, 0.88]; see Fig. 2 for the survival curve by exposure to cognitive techniques). This ratio was constant over time (Wald  $\chi^2$  [1] = 3.30,  $p = .069$ ). In contrast, fixed 2-year violent recidivism rates did not significantly differ as a function of cognitive technique exposure (12.1% for those exposed to cognitive techniques vs. 10.4% for those not exposed to cognitive techniques; odds ratio = 1.18 [0.79, 1.75]) and the hazard ratio indicated no difference in survival curves controlling for varying follow-up times (hazard ratio = 1.03 [0.76, 1.39]).

## 6. Discussion

The implementation of STICS in BC community corrections had two intended outcomes. First, there was the expectation that PO behavior would change and more closely align to RNR-based practices. Second, if this first outcome materialised, then client reconviction rates were expected to decrease with exposure to interventions from STICS-trained officers. Furthermore, the roll-out of STICS across the province not only afforded an opportunity to conceptually replicate earlier evaluations of the STICS model, but also to learn how well STICS can be adopted on a larger scale.

### 6.1. Enhancing community supervision through adherence to RNR

The dominant model of community supervision is the case management model (Bourgon, 2013). This model sees the PO as brokering community services for the client, with rather little attention paid to whether or not the services are based in evidence and effective in

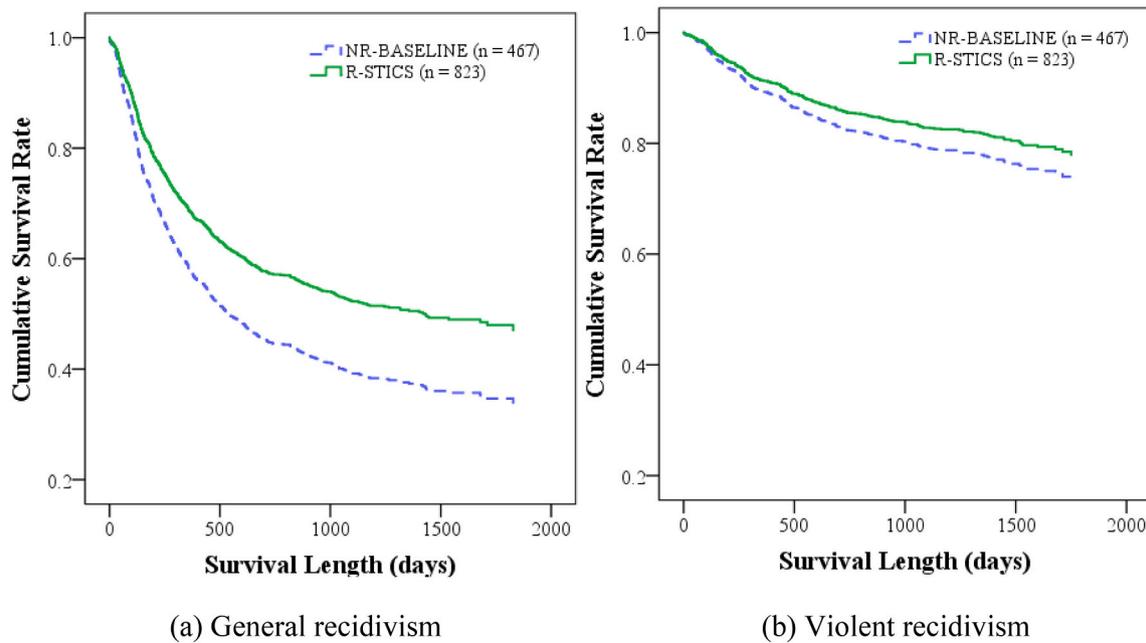


Fig. 1. Survival curves of NR-baseline and R-STICS groups for general and violent recidivism.

Note. The general and violent recidivism survival curves for each group were plotted after controlling for their age, criminal history, and time offence-free in the community before selection into the study.

Table 7

Cox regression predicting general recidivism from discussion of criminogenic needs.

	HR	95% CI
Step 1: Covariates		
R-Baseline (N = 281) vs. R-STICS (N = 820)	1.24	[0.996, 1.54]
Criminal history subscale (LSI-R)	<b>1.21</b>	[1.16, 1.26]
Age (year)	<b>0.98</b>	[0.97, 0.99]
Time offence-free <sup>a</sup> (month)	<b>0.97</b>	[0.95, 0.997]
Step 2: Discussion of criminogenic needs <sup>b</sup>		
	$\Delta\chi^2$ (df = 5) = 15.40, p = .009	
Procriminal Attitudes	<b>0.95</b>	[0.91, 0.99]
Procriminal Peers	1.04	[0.99, 1.08]
Family/marital	1.01	[0.97, 1.05]
Employment/education	1.01	[0.97, 1.05]
Substance abuse	<b>1.04</b>	[1.01, 1.07]

Note. HR = hazard ratio. CI = confidence interval. Numbers in bold indicate statistical significance (p < .05).

$\Delta\chi^2$  indicates  $\chi^2$  change from Step 1 to Step 2.

<sup>a</sup> Month(s) between the effective supervision start date and the first audio recording date for R-STICS clients; <sup>b</sup>A unit of each variable is set to be 0.10 (i.e., 10%).

reducing recidivism. Rather than the POs playing an active, face-to-face “helper” role much of what POs do on a daily basis is write presentence reports and engage in other administrative tasks (Bourgon, Rugge, Chadwick, & Bonta, 2018). We know that the RNR model explains much regarding the effectiveness of offender treatment programs (Bonta & Andrews, 2017). This requires a shift in policies and practices where the priorities are promoting an active change agent role for POs rather than an enforcement role even going so far as hiring, training, and evaluating POs to achieve a greater alignment of community supervision with RNR (Bourgon, 2013; Bourgon, Chadwick, & Rugge, 2020). STICS is but one component of organizational change.

The present study demonstrated that with training, POs could shift their attention to what matters: devote a greater proportion of the supervision sessions to criminogenic needs (with a corresponding decrease in time spent on noncriminogenic needs and conditions of probation). According to the need principle, targeting criminogenic needs is

Table 8

Cox regression predicting general recidivism rates from intervention skills.

	HR	95% CI
Model 1: Cognitive Skills		
Step 1: Covariates		
R-Baseline (N = 280) vs. R-STICS (N = 822)	1.19	[0.95, 1.48]
Criminal history subscale (LSI-R)	<b>1.21</b>	[1.16, 1.26]
Age (year)	<b>0.98</b>	[0.97, 0.99]
Time offence-free <sup>b</sup> (month)	<b>0.97</b>	[0.95, 0.996]
Step 2: Cognitive Intervention Skills		
	$\Delta\chi^2$ (df = 1) = 5.28, p = .022	
Cognitive skills	<b>0.90</b>	[0.81, 0.99]
Model 2: All Intervention Skills		
Step 1: Covariates and Cognitive Intervention Skills		
R-Baseline (N = 280) vs. R-STICS (N = 822)	1.11	[0.88, 1.40]
Criminal history subscale (LSI-R)	<b>1.21</b>	[1.16, 1.27]
Age (year)	<b>0.98</b>	[0.97, 0.99]
Time offence-free <sup>a</sup> (month)	<b>0.97</b>	[0.95, 0.997]
Cognitive skills	0.90	[0.81, 1.01]
Step 2: Other Intervention Skills		
	$\Delta\chi^2$ (df = 3) = 5.84, p = .120	
Collaborative skills	1.07	[0.98, 1.17]
Structuring skills	1.13	[0.98, 1.30]
Behavior skills	0.93	[0.80, 1.09]

Note. HR = hazard ratio. CI = confidence interval.  $\Delta\chi^2$  indicates  $\chi^2$  change from Step 1 to Step 2. Scores for each intervention skill were standardised (i.e., z-scores). Numbers in bold indicate statistical significance (i.e., p < .05).

<sup>a</sup> Month(s) between the effective supervision start date for NR-Baseline and the first audio recording date for R-STICS clients.

associated with reduced recidivism (Bonta, 2019; Bonta & Andrews, 2017; Lipsey & Cullen, 2007). There is also some research suggesting that increasing time spent on probation conditions is associated with increased recidivism (Bonta et al., 2008).

Of particular importance was the increased attention to attitudes. The average proportion of each supervision session dealing with pro-criminal attitudes increased substantially from pre- to post-training.

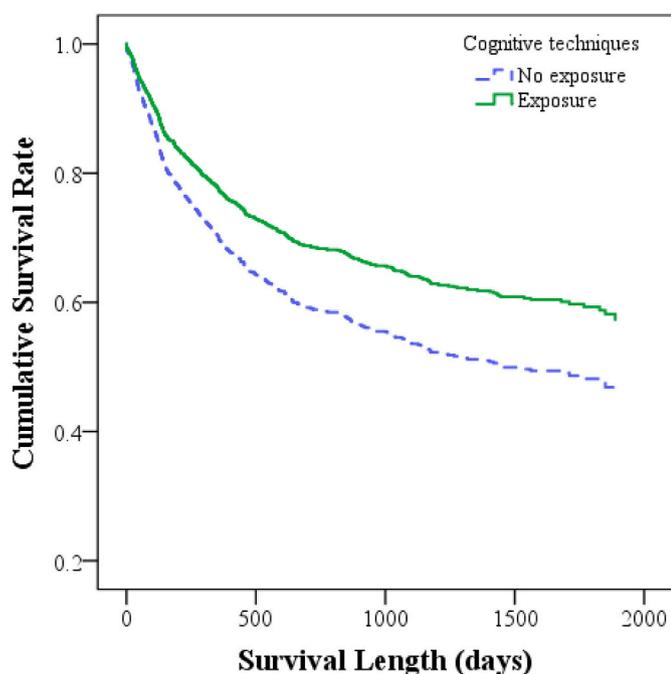


Fig. 2. General recidivism survival curves by exposure to cognitive intervention skills.

Note. Recidivism survival curves by exposure to cognitive intervention skills were estimated after accounting for client groups (R-Baseline vs. R-STICS), age, criminal history (LSI-R), and time offence-free.

This is consistent with earlier evaluations of the STICS model (Bonta et al., 2011; Bonta, Rugge, Bourgon, & Wanamaker, 2019), although in the current study, a larger proportion of each supervision session was spent discussing procriminal attitudes (0.18) compared to the 2011 (0.13) and 2019 (0.05) evaluations. It was surprising to find that the present results were greater than in the smaller demonstration evaluations, where effect sizes tend to be larger (Bonta & Andrews, 2017). The training and clinical support activities may have been more effective with respect to attitudes than the earlier studies, but we were unable to assess this possibility.

More time spent discussing procriminal attitudes necessarily leaves less time for discussions of other criminogenic needs. However, any discussion of attitudes *underlying* the other criminogenic needs in this study was coded as procriminal attitudes regardless of their content; therefore, it is possible that other criminogenic needs were still being discussed in the context of their underlying attitudes. Unfortunately, the coding methodology was not sensitive to this nuance in conversations between the PO and their client. Importantly, time spent on each criminogenic need was more balanced in post-training sessions than in pre-training sessions, which were skewed toward family/marital, employment/education, and substance abuse. This is an important shift as it indicates that more of the clients' criminogenic needs are being addressed in supervision sessions, particularly procriminal attitudes.

Adherence to the responsivity principle was reflected in the significant improvements following training across all four specific measures of cognitive-behavioral intervention skills (i.e., collaboration, structuring, cognitive intervention, and behavioral intervention). It was notable that this evaluation found a significant improvement in behavioral intervention skills (e.g., prosocial modeling, reinforcing prosocial behavior) whereas prior evaluations (Bonta et al., 2011; Bonta, Rugge, Bourgon, & Wanamaker, 2019) did not. Early work by Trotter (1996) had established the significance of prosocial modeling by POs to reduce client recidivism, with Dowden and Andrews' (2004) meta-analysis expanding the reach of behavioral intervention techniques to include effective reinforcement and disapproval. We speculate that the ongoing

clinical support activities in the current implementation may have enhanced the POs performance of behavioral skills. Research indicates that clinical support is important for the development of new skills (Beidas & Kendall, 2010; Bourgon et al., 2010). Across BC community corrections, POs were required to attend monthly meetings and expected to participate in other ongoing clinical support activities (in the 2019 study, clinical support activities were almost non-existent).

Given that STICS training resulted in improvements in PO behavior, their clients were expected to show reductions in recidivism. Indeed, R-STICS clients (audio-recorded clients supervised by STICS-trained POs) showed significantly lower 2-year general (43.0% vs. 61.4%) and violent (14.9% vs. 21.2%) recidivism rates compared to the NR-Baseline clients (randomly selected non-recorded baseline clients). Similarly, when the entire follow-up period for each client was considered, survival analyses showed that R-STICS clients had lower general recidivism rates after accounting for varying follow-up periods, age, risk, and time offence-free before selection into the study.

This is the first evaluation of STICS that measured violent recidivism as an outcome. In general, violent recidivism has a far lower base rate than general recidivism (Mossman, 1994). Therefore, larger offender samples are needed to detect violence. Previous evaluations of STICS had relatively small samples of probationers: 143 in the original evaluation (Bonta et al., 2011) and 123 in the 2019 evaluation (Bonta, Rugge, Bourgon, & Wanamaker, 2019). The current study had a much larger sample of 1194 probationers. Although the 2-year violent recidivism rate was significantly lower for R-STICS compared to NR-Baseline clients, the survival analyses found that group membership did not significantly predict violent recidivism (even though the hazard ratio suggested that R-STICS clients were on average 11% less likely to be reconvicted of a violent offence). It is possible that, because violent recidivism typically takes longer to occur than general recidivism, the 2-year follow-up period may not sufficiently reflect violent recidivism to evaluate the effects of STICS.

When specific PO behaviors were examined, actions consistent with RNR predicted reductions in general recidivism among R-STICS and R-Baseline clients. In line with the theory that attitudes underlie all other criminogenic needs (Bonta & Andrews, 2017), results suggest that procriminal attitudes may be the single most important criminogenic need associated with general recidivism. Consistent with Bonta et al. (2011), the use of cognitive intervention skills significantly predicted lower general recidivism rates in the current study – a finding that was not replicated by Bonta, Rugge, Bourgon, & Wanamaker, 2019. A possible explanation for the latter result is the very low proportion of each session devoted to attitudes in the 2019 evaluation, where an average of only 5% of the segments in each session focused on attitudes. The 2011 study showed that, on average, 13% of the segments in each session dealt with procriminal attitudes, whereas it was 18% in the present study. These findings suggest that there may be a certain dosage required for addressing attitudes to have an impact on client recidivism.

Contrary to expectation, a higher proportion of each session spent discussing substance abuse was associated with an *increase* in general recidivism. It is unclear why this would be the case, but we speculate that discussions surrounding substance abuse may be associated with a common condition of probation (e.g., submit to urinalysis). This surveillance/enforcement role may make a collaborative working relationship between a PO and their client more difficult (Skeem, Louden, Polaschek, & Camp, 2007; Trotter, 2013; Ugwu-dike & Raynor, 2013). Moreover, a greater focus on substance abuse leaves less time for discussions of procriminal attitudes, which appear to be particularly important for reducing general recidivism.

The responsivity principle states that cognitive-behavioral interventions are the most effective with offenders. Thus, teaching POs such interventional skills is an important goal of STICS. However, we have learned from previous evaluations of STICS and the current study that cognitive techniques, per se, may be an important therapeutic ingredient in contrast to purely behavioral interventions. In the Bonta

et al. (2011) evaluation, exposure to cognitive intervention techniques was associated with a 2-year general recidivism rate of 19%, while those not exposed to cognitive techniques had a rate of 37%. In the 2019 evaluation (Bonta, Rugge, Bourgon, & Wanamaker, 2019), there was no difference in the 2-year recidivism rate between the experimental and control clients, but survival analyses controlling for age and risk indicated a lower recidivism rate. In the current study, however, violent recidivism did not significantly differ as a function of cognitive intervention exposure, regardless of the follow-up period. This suggests that, for high-risk violent clients, more intense interventions (e.g., more frequent interventions and longer time) might be required to change their criminogenic needs in accordance with the risk principle.

Behavioral intervention skills were not significantly associated with reductions in general recidivism in this study or the 2011 evaluation of STICS. However, we would also argue that behavioral intervention techniques remain an integral part of the STICS model. Cognitive interventions may change a client's identity and personal narrative (Lloyd, Perley-Robertson, & Serin, 2020), while behavioral interventions (e.g., reinforcement, modeling) shape and consolidate the behaviors that affirm a personal narrative.

### 6.2. Large scale implementation: implications for policy and practice

Demonstration projects generally result in more substantial effects than large scale implementations, partly because they have fewer participants allowing the researchers to more closely monitor the evaluation (Bonta & Andrews, 2017). In addition, participants tend to be volunteers and possibly more conscientious in following the treatment and research protocols. Both the original STICS demonstration project (Bonta et al., 2011) and a subsequent demonstration project (Bonta, Rugge, Bourgon, & Wanamaker, 2019) had less than 100 POs, and all were volunteers. The current study involved approximately 350 POs, many of whom would not have volunteered for such a project and may have been resistant to change.

Our findings suggest that applying the intervention techniques developed in demonstration projects can be successfully transferred to a larger "real world" setting. There are a number of potential explanations that may account for the positive results found in the present study. First, the evaluators were closely involved in the delivery of STICS and design of the evaluation, an important aspect of positive treatment effects. For example, in a meta-analytic review of correctional treatment programs, Petrosino and Soydan (2005) found an average effect size of 0.16 (Cohen's *d*) associated with positive outcomes when the evaluators were involved in the development of the program vs. 0.02 when the evaluators were external to the program. One may take a cynical view and proclaim that the evaluators bias the results. However, it is just as likely that the commitment to the program by the evaluators ensures high fidelity. It is advantageous for correctional agencies to involve program developers as much as possible.

Second, the roll-out was closely monitored and, when it was determined that staff cooperation was below expectations, an integrity assessment was conducted. Specifically, at the mid-way point, the number of audio recordings submitted were below what was projected. The researchers and senior management then paused the implementation to conduct interviews with front line staff to explore the difficulties POs were experiencing and take corrective steps to improve adherence to the STICS model and evaluation (Bonta, Bourgon, Rugge, Gress, & Gutierrez, 2013). Without monitoring, fidelity to the protocols is threatened, and ensuring such monitoring is essential to success (Faringera, Durieza, Manchaka, & Sullivan, 2019). Organizations adopting new supervision models must be prepared to institute policies and practices to support and closely monitor implementation.

Third, considerable resources were allocated to the implementation (i.e., to permit POs to attend trainings, participate in clinical support activities, and even hiring additional staff). In other words, agencies must be prepared to adequately invest in a new model like STICS.

Fourth, implementation took three and one-half years (excluding the six-month pause in training). The extant evidence suggests that successful large-scale implementations take years (Bierman et al., 2002; Fixsen, Naoom, Blase, Friedman, & Wallace, 2005), and this slow approach in our case was advantageous.

### 6.3. Limitations

Because the implementation of STICS required all POs to be trained within a restricted timeframe, a randomised experiment to establish a causal relationship was not conducted. Instead, changes in PO behavior were evaluated by a pre-and post-test design which may be affected by history and maturation (Campbell & Stanley, 1966). However, the relatively short time frame POs had to complete all their audio recordings (approximately 12–15 months) minimizes history and maturation effects on the quality of PO behavior assessed prior to and after training.

The evaluation of client recidivism used a quasi-experimental design. This design is constrained by differences between comparison groups, in our case, age, risk, and time offence free before selection into the study. We attempted to minimize this threat to internal validity by controlling for these covariates. Furthermore, a randomly selected sample of non-recorded baseline clients (NR-Baseline) was used as a control group to reduce selection bias. Although NR-Baseline clients were likely more representative of the population of probationers before the implementation of STICS, there were still some limitations. For instance, all the R-STICS clients had agreed to have their supervision sessions recorded. Because we do not know whether NR-Baseline clients would have consented to have their supervision sessions audio recorded, it is possible that R-STICS and NR-Baseline clients differ in important ways for which we could not control (e.g., motivation to change, agreeableness).

A large portion of baseline clients (65.7% of R-Baseline and 40.3% of NR-Baseline) may have been exposed to STICS to some degree. That is, for some, their probation periods overlapped with their PO's training date. Ideally, baseline clients would not have been exposed to STICS-trained POs at all. Unfortunately, for the R-Baseline group, this was not possible given the system-wide roll-out of STICS. However, in hindsight, NR-Baseline clients who did not have probation orders that overlapped with their PO's training date could have been selected. On the one hand, selecting more recently supervised clients reduced history and maturation effects. On the other hand, this method increased contamination across groups, which may have attenuated the difference between R-STICS and R/NR-Baseline clients. In addition, the act of recording sessions after receiving training may have introduced participant bias/demand characteristics, such that STICS-trained POs would be on their "best behavior" for the recordings. Nonetheless, even if recordings reflect "best behavior" on the part of POs, these findings still speak to the program's efficacy (i.e., effectiveness of a program under ideal circumstances).

Finally, the coding of discussions of antisocial personality did not meet the threshold set for interrater reliability. Therefore, in the current study, we could not reliably examine the proportion of time spent discussing antisocial personality relative to other criminogenic needs. Antisocial personality is an important predictor of criminal recidivism (Bonta & Andrews, 2017), and therefore, important for POs to address. The audio coding instructions for antisocial personality should be revised for future research to improve the reliability of this variable. This would allow forthcoming studies to examine the potential impact of addressing antisocial personality in community supervision on client recidivism.

### 6.4. Summary and conclusions

Recognizing the aforementioned methodological limitations, the present evaluation provided concrete evidence for the value of the STICS

model. PO behavior changed in the expected direction with greater adherence to the need and responsivity principles. There were significant improvements with respect to POs' focus on problematic attitudes and cognitive-behavioral intervention skills. Additionally, probationers with greater exposure to STICS had significantly lower general recidivism rates than a random sample of probationers with less exposure to STICS. Addressing procriminal attitudes using cognitive intervention techniques stood out as being particularly important to reducing general recidivism. The importance of changing attitudes as the first step in psychotherapeutic change dates back at least 60 years: "the aims of all forms [of psychotherapy] are to help the patient correct the attitudes causing him trouble" (Frank, 1963, p. 41). Therefore, our results are consistent with the broader literature on offender rehabilitation and behavior change. In conclusion, the current study suggests that STICS may be a promising model for improving the effectiveness of community supervision.

#### Author note

The opinions, findings, conclusions, and recommendations expressed are those of the authors and do not necessarily represent the views of Public Safety Canada.

#### Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or non-for-profit sectors.

#### Ethical approval

The research was conducted in accordance with the ethical standards of the Province of British Columbia and BC Corrections and the Canadian Tri-council Policy Statement: Ethical Conduct for Research Involving Humans.

#### Informed consent

Informed written consent was obtained from all individual participants included in the study.

#### Declaration of Conflict of Interest

James Bonta, Guy Bourgon, and Tanya Ruge are co-authors of STICS. The copyright for STICS is held by the Government of Canada, and none of the authors receive royalties from this program. James Bonta receives royalties on sales of the Level of Service Inventory-Revised cited in this article.

#### Acknowledgments

We would like to thank Bill Small, Carrie McCulley, the probation officers, and their managers, for their efforts in meeting the demands of our research protocols. At Public Safety Canada our gratitude is extended to Leticia Gutierrez who trained and supervised the research assistants who coded the audio recordings and, of course, the research assistants themselves. Finally, we thank the trainers and the senior managers for their support in this project.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jcrimjus.2021.101816>.

#### References

- Andrews, D. A., & Bonta, J. (1995). *The level of service inventory – Revised*. Toronto: Multi-Health Systems.
- Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, 17(1), 19–52. <https://doi.org/10.1177/0093854890017001004>.
- Andrews, D. A., & Dowden, C. (2006). Risk principle of case classification in correctional treatment: A meta-analytic investigation. *International Journal of Offender Therapy and Comparative Criminology*, 50(1), 88–100. <https://doi.org/10.1177/0306624X05282556>.
- Beidas, R. S., & Kendall, P. C. (2010). Training therapists in evidence-based practice: A critical review of studies from a systems-contextual perspective. *Clinical Psychology: Science and Practice*, 17(1), 1–30. <https://doi.org/10.1111/j.1468-2850.2009.01187.x>.
- Bierman, K. L., Coie, J. D., Dodge, K. A., Greenberg, M. T., Lochman, J. E., McMahon, R. J., & Pinderhughes, E. (2002). The implementation of the Fast Track Program: An example of a large-scale prevention science efficacy trial. *Journal of Abnormal Child Psychology*, 30(1), 1–17. <https://doi.org/10.1023/A:1014292830216>.
- Bonta, J. (2019). Criminogenic needs. In R. D. Morgan (Ed.), *The SAGE encyclopedia of criminal psychology* (pp. 324–328). Sage Publications. <https://doi.org/10.4135/9781483392240.n372>.
- Bonta, J., & Andrews, D. A. (2017). *The psychology of criminal conduct* (6th ed.). Routledge.
- Bonta, J., Bourgon, G., Ruge, T., Gress, C., & Gutierrez, L. (2013). Taking the leap: From pilot project to wide-scale implementation of the Strategic Training Initiative in Community Supervision (STICS). *Justice Research and Policy*, 15(1), 17–35. <https://doi.org/10.3818/JRP.15.1.2013.17>.
- Bonta, J., Bourgon, G., Ruge, T., Scott, T. L., Yessine, A. K., Gutierrez, L., & Li, J. (2011). An experimental demonstration of training probation officers in evidence-based community supervision. *Criminal Justice and Behavior*, 38(11), 1127–1148. <https://doi.org/10.1177/0093854811420678>.
- Bonta, J., Ruge, T., Bourgon, G., & Wanamaker, K. (2019). A conceptual replication of the Strategic Training Initiative in Community Supervision (STICS). *Journal of Experimental Criminology*, 15(3), 397–419. <https://doi.org/10.1007/s11292-019-09371-4>.
- Bonta, J., Ruge, T., Scott, T., Bourgon, G., & Yessine, A. (2008). Exploring the black box of community supervision. *Journal of Offender Rehabilitation*, 47(3), 248–270. <https://doi.org/10.1080/10509670802134085>.
- Bonta, J., Wallace-Capretta, S., & Rooney, J. (2000). A quasi-experimental evaluation of an intensive rehabilitation supervision program. *Criminal Justice and Behavior*, 27(3), 312–329.
- Bourgon, G. (2013). The demands on probation officers in the evolution of evidence-based practice: The forgotten foot soldiers of community corrections. *Federal Probation*, 77(2), 30–34.
- Bourgon, G., Bonta, J., Ruge, T., & Gutierrez, L. (2010). Technology transfer: The importance of on-going clinical supervision in translating "What Works" to everyday community supervision. In F. McNeil, P. Raynor, & C. Trotter (Eds.), *Offender supervision: New directions in theory, research and practice* (pp. 88–106). Willan Publishing.
- Bourgon, G., Chadwick, N., & Ruge, T. (2020). Beyond core correctional practice: Facilitating prosocial change through the Strategic Training Initiative in Community Supervision. In J. S. Wormith, L. A. Craig, & T. Hogue (Eds.), *Handbook of what works in violence risk management: Theory, research and practice* (pp. 505–526). John Wiley and Sons.
- Bourgon, G., Ruge, T., Chadwick, N., & Bonta, J. (2018). The living laboratory studies: Providing insights into community supervision practices. *Federal Probation*, 82(1), 3–12.
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Rand McNally College Publishing.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284–290. <https://doi.org/10.1037/1040-3590.6.4.284>.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159. <https://doi.org/10.1037//0033-2909.112.1.155>.
- Cullen, F. T. (2013). Rehabilitation: Beyond nothing works. In M. Tonry (Ed.), *Crime and justice, Vol. 42, Crime and justice in America 1975–2025* (pp. 299–376). University of Chicago Press.
- Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, 25(1), 7–29. <https://doi.org/10.1177/0956797613504966>.
- Dowden, C., & Andrews, D. A. (2000). Effective correctional treatment and violent reoffending: A meta-analysis. *Canadian Journal of Criminology and Criminal Justice*, 42(4), 449–467. <https://doi.org/10.3138/cjcrim.42.4.449>.
- Dowden, C., & Andrews, D. A. (2004). The importance of staff practices in delivering effective correctional treatment: A meta-analysis of core correctional practices. *International Journal of Offender Therapy and Comparative Criminology*, 48(2), 203–214. <https://doi.org/10.1177/0306624X03257765>.
- Farrington, A. J., Durieza, S. A., Manchaka, S. M., & Sullivan, C. C. (2019). Adherence to "What Works": Examining trends across 14 years of correctional program assessment. In *Corrections: Policy, practice and research*. <https://doi.org/10.1080/23774657.2019.1659193>. Advance online publication.
- Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231).

- Flores, A. W., Holsinger, A. M., Lowenkamp, C. T., & Cohen, T. H. (2017). Time-free effects in predicting recidivism using both fixed and variable follow-up periods: Do different methods produce different results. *Criminal Justice and Behavior*, 44(1), 121–137. <https://doi.org/10.1177/0093854816678649>.
- Frank, J. D. (1963). *Persuasion and healing: A comparative study of psychotherapy*. Schocken Books.
- Gendreau, P., & Smith, P. (2007). Influencing the “people who count”: Some perspectives on the reporting of meta-analytic results for prediction and treatment outcomes with offenders. *Criminal Justice and Behavior*, 34(12), 1536–1559. <https://doi.org/10.1177/0093854807307025>.
- Hanson, R. K. (2002). Recidivism and age: Follow-up data from 4,673 sexual offenders. *Journal of Interpersonal Violence*, 17(10), 1046–1062. <https://doi.org/10.1177/088626002236659>.
- Hanson, R. K., Harris, A. J. R., Letourneau, E., Helmus, L. M., & Thornton, D. (2018). Reductions in risk based on time offense-free in the community: Once a sexual offender, not always a sexual offender. *Psychology, Public Policy, and Law*, 24(1), 48–63. <https://doi.org/10.1037/law0000135>.
- Hirschi, T., & Gottfredson, M. R. (1993). Age and the explanation of crime. *American Journal of Sociology*, 89(3), 552–584. <https://doi.org/10.1086/227905>.
- Kaeble, D., & Alper, M. (2020). *Probation and parole in the United States, 2017–2018*. Washington: Bureau of Justice Statistics.
- Labrecque, R. M., Smith, P., Schweitzer, M., & Thompson, C. (2013). Targeting antisocial attitudes in community supervision using the EPICS model: An examination of change scores on the Criminal Sentiment Scale. *Federal Probation*, 77(3), 15–20. [https://www.uscourts.gov/sites/default/files/77\\_3\\_2\\_0.pdf](https://www.uscourts.gov/sites/default/files/77_3_2_0.pdf).
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310U5iVc5tA6im->
- Lipsey, M. W., & Cullen, F. T. (2007). The effectiveness of correctional rehabilitation: A review of systematic reviews. *Annual Review of Law and Social Science*, 3(1), 297–320. <https://doi.org/10.1146/annurev.lawsocsci.3.081806.112833>.
- Lloyd, C. D., Perley-Robertson, B., & Serin, R. C. (2020). Age and strengths in a community corrections sample. *International Journal of Forensic Mental Health*, 19(3), 215–226. <https://doi.org/10.1080/14999013.2019.1684404>.
- Lowenkamp, C. T., Latessa, E. J., & Holsinger, A. M. (2006). The risk principle in action: What have we learned from 13,676 offenders and 97 correctional programs. *Crime & Delinquency*, 52(1), 77–93. <https://doi.org/10.1177/001128705281747>.
- Malakieh, J. (2020). Adult and youth correctional statistics in Canada, 2018/2019. <https://www150.statcan.gc.ca/n1/pub/85-002-x/2020001/article/00016-eng.pdf>.
- Martinson, R. (1974). What works?—Questions and answers about prison reform. *The Public Interest*, 35, 22–54.
- Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology*, 62(4), 783–792. <https://doi.org/10.1037//0022-006x.62.4.783>.
- Petrosino, A., & Soydan, H. (2005). The impact of program developers as evaluators on criminal recidivism: Results from meta-analyses of experimental and quasi-experimental research. *Journal of Experimental Criminology*, 1(4), 435–450. <https://doi.org/10.1007/s11292-005-3540-8>.
- Robinson, C. J., Lowenkamp, C. T., Holsinger, A. M., VanBenschoten, S., Alexander, M., & Oleson, J. C. (2012). A random study of staff training aimed at reducing re-arrest (STARR): Using core correctional practice in probation interactions. *Journal of Crime and Justice*, 35(2), 167–188. <https://doi.org/10.1080/0735648X.2012.674823>.
- Skeem, J. L., Loudon, J. E., Polaschek, D., & Camp, J. (2007). Assessing relationship quality in mandated community treatment: Blending care with control. *Psychological Assessment*, 19(4), 397–410. <https://doi.org/10.1037/1040-3590.19.4.397>.
- Smith, A., Heyes, K., Fox, C., Harrison, J., Kiss, Z., & Bradbury, R. (2018). The effectiveness of probation supervision towards reducing reoffending: A rapid evidence assessment. *Probation Journal*, 65(4), 407–428. <https://doi.org/10.1177/0264550518796275>.
- Trotter, C. (1996). The impact of different supervision practices in community corrections: Cause for optimism. *Australian & New Zealand Journal of Criminology*, 29(1), 1–19. <https://doi.org/10.1177/000486589602900103>.
- Trotter, C. (2013). Reducing recidivism through probation supervision: What we know and don't know through four decades of research. *Federal Probation*, 77(2), 43–48.
- Ugwudike, P., & Raynor, P. (2013). *What works in offender compliance: International perspectives and evidence-based practice*. Palgrave Macmillan.
- Wormith, J. S., & Zidenberg, A. M. (2018). The historical roots, current status, and future applications of the Risk-Need-Responsivity model (RNR). In E. L. Jeglic, & C. Calkins (Eds.), *New frontiers in offender treatment* (pp. 11–41). [https://doi.org/10.1007/978-3-030-01030-0\\_2](https://doi.org/10.1007/978-3-030-01030-0_2). Springer Nature Switzerland AG.