The Intensive Support and Supervision Program: A Best Practice File Evaluation of Implementation and Outcomes

by

Sarah Lindsay Mordell

M.A., Simon Fraser University, 2009

Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Psychology
Faculty of Arts and Social Sciences

SIMON FRASER UNIVERSITY

Summer 2014

Approval

Name: **Sarah Lindsay Mordell**

Degree: **Doctor of Philosophy**

Title: ***The Intensive Support and Supervision Program: A Best Practice File Evaluation of Implementation and Outcomes***

Examining Committee: **Chair:** Wendy Thornton
Associate Professor

Ronald Roesch
Senior Supervisor
Professor

Jodi Viljoen
Supervisor
Associate Professor

Stephen Hart
Supervisor
Professor

William Glackman
Internal Examiner
Associate Professor
School of Criminology

Joel Dvoskin
External Examiner
Assistant Professor
Department of Psychiatry
University of Arizona

Date Defended/Approved: August 11, 2014

Partial Copyright Licence

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

c. as a co-investigator, collaborator or research assistant in a research project approved in advance,

or

d. as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

update Spring 2010

Abstract

The Youth Justice Intensive Support and Supervision Program (ISSP) is offered to high-risk, justice-involved youth as an alternative to custody or as a means of facilitating community re-entry after a period of custody. The aim of ISSP is to reduce recidivism by increasing supervision and supporting activities to reduce participants' criminogenic needs. Research on similar programs has yielded mixed findings for reductions in recidivism, although the way programs are delivered appears to be a key factor influencing their effectiveness.

For the current study, an evidence-based evaluation framework of best practices was developed for ISSP from the research literature on similar and general youth justice programs. File data for 176 ISSP participants were used to evaluate the program's delivery against the best practice framework as well as ISSP's program guidelines to determine how well the program was implemented and whether a better-implemented program led to better outcomes. A questionnaire study was also conducted with Youth Probation Officers.

Adherence to individual program guidelines and best practices varied from 11% to 92% of youth, while the mean of overall best practices implemented was approximately 50%. An implementation composite of best practices was associated with marginal reductions in multiple recidivism outcomes and the positive relationships with intermediate targets such as school and employment outcomes approached significance. Moreover, the relationship between ISSP implementation and recidivism was strongest for younger, for Aboriginal, and for higher-risk participants. An index of level of engagement in ISSP also predicted reductions in recidivism in the year following ISSP. The quality of the ISSP documentation on the files limited the interpretation of the study findings. Several recommendations are offered to support and improve current practice.

Keywords: Juvenile justice; intervention; evidence-based; best practices; Intensive Support and Supervision Program

Acknowledgements

This project and dissertation was supported by the efforts of a great many individuals, for whom a brief acknowledgement hardly seems sufficient. Nevertheless, I am deeply indebted to my senior supervisor, Dr. Ronald Roesch, who has been an incredible mentor and has shaped my graduate career in countless ways. He along with Dr. Jodi Viljoen provided opportunities to work with British Columbia Youth Justice early in my studies, which inspired this research. I am grateful also to the feedback provided by my committee members Dr. Viljoen and Dr. Stephen Hart at various stages of the project's development. I would furthermore like to thank my external examiners Dr. Joel Dvoskin and Dr. William Glackman, whose valuable suggestions and comments I will greatly benefit from as I move forward with this research.

This study required considerable administrative support and would not have been possible without the generous assistance of British Columbia Youth Justice. In particular, Phil Peachy, Chris Zatylny, and Nerina Holderness contributed notable time and resources, but there were several others who provided training, access, and review. Numerous youth probation administrative staff and officers provided files and space for to review them with great efficiency and courtesy. British Columbia Corrections allowed for the collection of adult recidivism data, which was essential to the validity of the study. I am also grateful to Gary Kumka and the Cambie team at Youth Forensic Psychiatric Services for providing space for the study and for their moral support. An extra special thank you is due to Elinor Craig who accepted and stored boxloads of files on my behalf. I furthermore appreciate the financial support provided by the Social Sciences Research and Humanities Council and an American Psychology-Law Society grant.

Regarding other project support, I would like to thank Billie Joe Rogers for completing the inter-rater reliability for the program coding and providing other invaluable feedback. Natalie Harrison and Sara Harvey also collected inter-rater and file data. I am additionally grateful to Dr. Kim van der Woerd and Nicole Muir who, along with Billie Joe, formed the ad-hoc Aboriginal Advisory Committee and contributed their time to provide interpretations of the study findings. Patricia Wallis supplied helpful statistical consultation.

Finally, since a dissertation does not occur in a social and emotional vacuum, I cannot thank my husband, Dharm, enough for his various forms of support and encouragement throughout this whole process. He has made many sacrifices along the way and never once asked “are we there yet?”. My family and extended family have always supported my decision to spend the better part of my young adulthood in school. Thanks to my sister Devon for providing much needed diversions along the way. I am especially glad that my 97-year-old grandmother, Mari, hung in there long enough to see me defend, among other things I am sure. I have enjoyed the many warm friendships with my fellow SFUers, particularly the cohort support. It has been an amazing journey!

Table of Contents

List of Tables

List of Figures

List of Acronyms

ANCOVA	Analysis of Covariance
CI	Confidence Interval
CORNET	Corrections Network
FASD	Fetal Alcohol Spectrum Disorder
ICC	Intra-class Correlation Coefficient
IRA	Inter-rater Agreement
IRR	Inter-rater Reliability
ISSO	Intensive Support and Supervision Order
ISSP	Intensive Support and Supervision Program
MAR	Missing at Random
MCAR	Missing Completely at Random
MH	Mental Health
MNAR	Missing Not at Random
RNR	Risk-Needs-Responsivity model
SAVRY	Structured Assessment of Violence Risk in Youth
SPEP	Standardized Program Evaluation Protocol
YCJA	*Youth Criminal Justice Act*
YCRNA	Youth Community Risk/Needs Assessment
YLS/CMI	Youth Level of Service/Case Management Inventory
YPO	Youth Probation Officer

Chapter 1.

Introduction

An Impetus for Non-custodial Alternatives

Under the *Youth Criminal Justice Act* (2002), changes were made to youth justice legislation to reduce the previous "over-reliance on incarceration for non-violent young persons" (YCJA, 2002; Preamble). These changes appear to reflect an understanding that custodial sentences are not a desirable outcome for youth. There are, in fact, several rationales for the youth criminal justice system to minimize the use of custodial measures. Of particular interest to taxpayers, incarcerating offenders is very expensive. The cost per adolescent to be housed in custody for one year is $215,000, as compared to $20,000 for offenders under community supervision (Representative for Children and Youth & Office of the Provincial Health Officer, 2009). The cost to keep youth under community supervision includes the cost of receiving specialized services like residential addictions treatment and psychological assessment and treatment through government agencies, such that it reflects more than just supervision of probation conditions.

More importantly, however, the costs of custody to the youth themselves are high. Though youth undoubtedly enter custody with significant problems in a number of areas, the experience of custody does not provide the ideal circumstances for rehabilitation. And while imprisonment is unlikely to benefit most, the use of custody may be particularly harmful for adolescents. Youth are undergoing important periods of psychosocial development, which may be disrupted by the experience of custody (Chung, Little, & Steinberg, 2005). The opportunities to develop maturity, independence, and prosocial values are limited in custody and offenders may find themselves lagging even further behind their same-age peers when seeking employment or forming relationships upon their release (Altschuler & Brash, 2004). With even fewer

opportunities for prosocial engagement upon leaving custody, it is likely that youth will find themselves falling back into criminogenic lifestyles.

Unsurprisingly, youth receiving custody sentences do not tend to have positive recidivism outcomes. Though Canadian data are not available, re-offence rates in the range of 40% to up to 85% for follow-up periods of between two and nine years have been found for youth released from custody in the United States (Bullis, Yovanoff, Mueller, & Havel, 2002; Trulson, Marquart, Mullings, & Caeti, 2005; van Marle, Hempel, & Buck, 2010). However, due to the at-risk nature of the sample, it is difficult to disentangle the contributions of pre-existing criminogenic needs and the experience of custody to recidivism outcomes.

Whether or not custody might contribute to recidivism in of itself, confining a youth to custody does not appear to convey any benefits over community alternatives in terms of deterring youth from offending. Due to developmental factors that impact youths' decision-making, the threat of custody is unlikely to dissuade youth from engaging in criminal behaviour (Steinberg, 2009). Systematic reviews of the research literature consistently show that youth who are incarcerated or given custodial sentences have the same or higher recidivism rates than those receiving non-custodial or community sentences (Aos, Phipps, Barnoski, & Lieb, 2001; Lipsey & Cullen, 2007). Within those sentenced to custody, a study that carefully controlled for selection biases found that longer sentences did not convey any additional reductions in recidivism for sentences between three and 13 months (Loughran et al., 2009). It would thus appear that custody merely restricts the opportunity for youth to offend for the brief time that they are in custody, which in Canada is a median length of 35 days (excluding time spent in custody prior to sentencing; Statistics Canada, n.d.).

Given the high costs and questionable effectiveness of custody in protecting the public and deterring further criminal behaviour, it is clearly desirable to avoid or reduce the use of custody for youth where feasible. Nonetheless, in the interest of public safety, it is important that high-risk youth are still subject to more intensive supervision than that provided by probation. Moreover, protection of the public in the long run will be best achieved by offering programming to youth to address their criminogenic needs. In order to accomplish these objectives, programs such as the Intensive Support and Supervision Program have emerged.

British Columbia's Youth Justice Intensive Support and Supervision Program

British Columbia has one of the lowest rates of youth in custody in Canada; reporting an average of 105 youth in custody during 2011-2012, which represents a 75% reduction from 1995-1996 (Ministry of Children and Family Development, 2012). These low numbers are credited to the changes in the *YCJA* (2002) from previous legislation and the use of such community-based alternative programs as the Youth Justice Intensive Support and Supervision Program (ISSP).

British Columbia's Youth Justice ISSP is based on the premise that increased surveillance and enforcement is not sufficient to prevent recidivism in the long term; the program also aims to address youths' needs through supportive services (Ministry of Children and Family Development, 2006). Youth are selected for the program based on their status as a high-risk offender and they are referred in order of priority according to sentence, offence, and risk characteristics. Adolescents referred to Youth Justice ISSP are assigned an ISSP worker or workers, who work one-on-one or in teams with youth to achieve goals related to their needs. ISSP workers are responsible for preparing a detailed service plan of goals and activities, monitoring probation conditions, supporting youths' participation in rehabilitative programs, liaising with parents and caregivers, and facilitating youths' participation in prosocial activities, among other duties.

The program sets out specific guidelines in terms of the quality and nature of contacts that ISSP workers should have with youth. Caseloads are limited to between six and 10 youth. There is some flexibility for individual youth and the amount of contact may vary based on youths' needs and their stage in the program (e.g., youth may require additional contacts at the beginning of the program that will later taper off). However, ISSP workers are expected to dedicate a minimum of two hours per week to each youth, including two in-person contacts, two telephone contacts, one contact with the parent or caregiver, and one or more contacts with community resources. Cases are re-evaluated every six months to determine whether youth should remain in the program (e.g., because of an ongoing high level of need).

However, the guidelines for Youth Justice ISSP do not provide much structure for the day-to-day operation of the program, which might include critical areas of focus for each youth and specific types of activities to address youths' needs. ISSP's flexibility to

select among priorities and activities is a significant strength for ISSP workers who have the training and background to benefit from it. Meanwhile, other ISSP workers may struggle to meet the goals of the program from the limited guidance and a lack of structure is likely to create great variability in the quality of the program's delivery. Other similar programs have noted the difficulty of translating theory or ideals into specific practice by staff and how this may lead to inconsistency in the implementation of the program (e.g., Lane, Turner, Fain, & Seghal, 2005).

Furthermore, there is a lack of outcome research on Youth Justice's ISSP to determine whether the program is actually achieving its aims. While ISSP is believed to reduce reoffending, some youth justice programs have been found to have no effect or even increase recidivism when evaluated (see, for example, meta-analyses on prison visitation programs or boot camps; Lipsey, 2009; Petrosino, Turpin-Petrosino, Hollis-Peel, & Lavenberg, 2013). Research on other programs with similar features to ISSP can provide an indication of whether the program might be expected to reduce recidivism and what components of the program are critical to this aim.

The State of the Literature on Similar Programs

Programs with similar aims to Youth Justice ISSP have been widely adopted internationally, with programs operating in the United States, Australia, and Britain. In North America, intensive supervision programs originated with adult offenders as early as the 1980s, but a treatment focus was mostly lacking (Paparozzi & Gendreau, 2005). Fortunately, youth programs have tended to focus more on rehabilitation.

Intensive Support and Supervision Program – Canada. Programs in Canada provide the most relevant comparison to Youth Justice ISSP in terms of the population and program factors since ISSP is a federal initiative (*YCJA*, 2002; Sec. 41(2)(l)). However, even at the most basic level, intensive support and supervision sentences appear to be meted out quite differently among the provinces by the Courts. Of the 471 total offences receiving an Intensive Support and Supervision Order (ISSO) in Canada in 2009-2010, 375 were sentenced to an ISSO in British Columbia (Statistics Canada, n.d.), while only 16 offences received an ISSO sentence in Ontario. While the number of community providers offering intensive support and supervision programs in Ontario identified in an Internet search would suggest that referrals more frequently come from

probation officers, these disparities nonetheless suggest significant differences in the way youth may be identified for intensive support and supervision programs, provincial support for the initiative, and how the program is delivered across the country.

Further hindering comparison, information on other ISSPs offered in Canada is limited to unpublished or "grey" literature and there is minimal evidence of formal evaluation of these programs. From the information available, it appears that the Ontario government provides an Intensive Support and Supervision Program to approximately 120 youth per year through multiple programs (Ministry of Children and Youth Services, 2011). A mental health diagnosis is necessary and the focus of the program is on mental health needs as opposed to criminogenic needs. The city of Calgary, Alberta, also provides an ISSO Initiative to a group of approximately 15 youth aged 15-24 under specific criteria (at least five previous findings of guilt, at least five convictions for failure to comply, a mental health diagnosis or indications thereof, minimal engagement in professional supports, and minimal involvement in productive activities; Costello & MacCrae-Krisa, 2011). It is therefore clear that even if evaluation research were available for programs offered by other provinces, the variability in how the program is structured and delivered would make it difficult to know whether program effectiveness would generalize across different provincial intensive support and supervision programs.

Within British Columbia, community providers are contracted to provide ISSP to youth who live outside the custody centre service regions or who are otherwise deemed more appropriate for community ISSP (e.g., lower-risk youth). One such organization is PLEA Community Services Society of British Columbia, for which limited outcome information is available for ISSP. Results from 18 interviews at three months into the program indicated that, while some youth reported improvement in their criminal justice involvement, as many or more youth reported no change and some reported decline (Vancouver Coastal Youth Justice Services, 2008). Improvements were further attenuated at six months. A later evaluation conducted by the McCreary Centre Society (2012) found declines in criminal charges, custody time, and risk behaviour for youth enrolled in PLEA programming. However, the evaluation covered multiple programs offered by PLEA, did not include a control group, and did not separate out the effects of ISSP or provide specific information on the activities of PLEA's ISSP.

International programs. Large, rigorous evaluation studies of comparable community re-entry programs for youth in custody and probation programs in the United States and the United Kingdom have generally found no differences between program and control youth for recidivism (Intensive Aftercare Program: Wiebush, Wagner, McNulty, Wang, & Le, 2005; Intensive probation: Lane et al., 2005; Intensive Supervision and Surveillance Programme: Gray, Taylor, Roberts, Merrington, Fernandez, & Moore, 2005; Serious and Violent Offender Re-entry Initiative: Hawkins, Dawes, Lattimore, & Visher, 2009; Lattimore & Visher, 2010; see also: Bouffard & Bergseth, 2008; Greenwood, Deschenes, & Adams, 1993; Iutcovich & Pratt, 1998; Land, McCall, & Williams, 1990). Some researchers have even gone so far as to recommend that the ISSP model be rejected altogether (e.g., Ellis, Pamment, & Lewis, 2009).

However, there are a number of reasons not to dismiss the ISSP model outright. Meta-analyses of intensive re-entry and probation programs have found small mean reductions in recidivism rates relative to youth receiving regular probation (Aos et al., 2001; James, Stams, Asscher, De Roo, & van der Laan, 2013; Lipsey, 1999). The reason that these effect sizes were too small to detect in individual (though large) studies may be due to inconsistency in the quality of program delivery. Wiebush and colleagues (2005) found that all of the Intensive Aftercare Program evaluation sites had at least one weak area of delivery, some of which were considered critical to program success. Also, the ratios of program service hours relative to the control group in custody and in the community varied considerably across sites. Similarly, an evaluation of the United Kingdom program found inconsistency in the program approaches across the 41 sites included (Moore, Gray, Roberts, & Taylor, 2004).

The program designs of these intensive re-entry or probation programs take the form of full days of structured activity to slightly more service than is offered by regular probation, even within the same program. Due to the variability within the programs as well as the variation in the programs' design from Youth Justice's ISSP, it is difficult to determine whether ISSP is likely to reduce recidivism from these evaluation data. Nonetheless, the likelihood to which ISSP will be effective in reducing recidivism is increased to the extent that the program principles align with research evidence.

Evidence-based Best Practice Guidelines for ISSP

Lipsey and colleagues offer a definition of evidence-based practice as "the effectiveness of the treatments, services, and programs provided to those in need should have been demonstrated in credible research prior to widespread use" (p.15; Lipsey, Howell, Kelly, Chapman, & Carver, 2010). To this end, they offer three possible means of achieving evidence-based practice available to juvenile justice agencies.

One evidence-based option would be to discard ISSP in favour of a model program that has been validated through rigorous research methods across multiple sites to reduce recidivism (see the Blueprints for Violence Prevention: Mihalic, Fagan, Irwin, Ballard, & Elliott, 2004). However, financial and staff resources for such programs are not always available and certainly not in the short-term. A sense of ownership has been identified as a key factor to program success, which may be difficult to maintain when an existing program is transported to a new site and could lead to non-evidence based modifications by staff (Mihalic, Irwin, Fagan, Ballard, & Elliott, 2004). Moreover, the effectiveness of model programs may be attenuated when offered to the more heterogeneous populations of real-world settings and without the service infrastructure or close supervision provided by the program developer (Lipsey et al., 2010). In fact, a null effect may be yielded for a model program when transported to a site with different characteristics from the evaluation sites, as was the case in a randomized control evaluation trial of the model program Multisystemic Therapy in Ontario (Cunningham, 2002). The evaluation indicated that the program did not yield improved outcomes over the control group, despite the oversight of the program developers' company during the first two years of the trial at a cost of over $200,000. The author concluded that the failure to find differences may have been due to an existing higher level of service (e.g., health care, community services) in Ontario for the control group relative to the model program evaluation studies' control groups in the United States.

A second possible means of achieving evidence-based practice is to subject the program to its own rigorous and ideally ongoing evaluation, which would provide the best evidence about the specific program's effectiveness as well as identify problems with implementation and subgroups for which the program proves to be the most beneficial (Lipsey et al., 2010). Unfortunately, given that a control group is generally a requirement for a rigorous research design, it is not ethically feasible to maintain this level of

evaluation. Other disadvantages include the resources and expertise required for evaluation, and the possibility that the findings of the study will no longer apply if there are changes to the population or program (Lipsey et al., 2010).

The third approach to evidence-based practice suggested by Lipsey and colleagues (2010) is to choose a general type of program (e.g., intensive probation) that has been shown to reduce recidivism through meta-analytic research and to deliver the program according to best practices drawn from juvenile justice research. Although there may be variability in the programs' elements, meta-analytic techniques can be used to code these features and determine which are significantly related to positive effect sizes. Since meta-analyses include studies spanning many settings and populations, best practices that emerge on average as positive are more robust in many ways than model programs (Lipsey et al., 2010). As an example of how these findings might be used in practice, Lipsey and colleagues (2010) proposed the Standardized Program Evaluation Protocol (SPEP) featuring six criteria derived from the meta-analytic research literature to which juvenile justice programs might be compared to determine the extent to which the program might be expected to reduce recidivism. A broad framework of 17 effective treatment principles for youth and adult offenders identified through a parallel process has also been proposed by Bonta and Andrews (Risk-Needs-Responsivity [RNR] model for offender assessment and rehabilitation; 2007).

Nevertheless, simply evaluating ISSP's design against the SPEP guidelines or the RNR model should not mark the end of this evidence-based approach. There is sufficient research evidence on similar programs to ISSP that a more specific set of best practices to ISSP could better inform evaluation. Furthermore, the SPEP guidelines, and to some extent the RNR model, are directed at a program level of evaluation; that is, whether the program's design is evidence-based. Fidelity of the program's design to the SPEP guidelines or other best practices is an important first step but provides little information about variation in the program's delivery across individuals. Also, evaluation at the program level may underestimate the true effects of best practices since it is generally based on the program's intended rather than actual delivery.

In all, using research evidence to guide the selection of specific best practices for ISSP and making revisions to the existing program guidelines if necessary is currently the most feasible option for evidence-based practice for Youth Justice ISSP. However,

even evidence-based practices need to be evaluated, as has been done with model programs and the SPEP guidelines, to ensure that these best practices continue to lead to reductions in recidivism in the program and population of interest. For example, an adult re-entry program that was designed according to empirical principles derived from the research literature resulted in worse outcomes for program participants on multiple measures of recidivism (Wilson & Davis, 2006). It should be noted, however, that the program focused somewhat narrowly on certain best practices while failing to address others, suggesting a need for a comprehensive best practice model for ISSP.

A framework for best practices. Prior to selecting best practices against which to evaluate Youth Justice ISSP, it is beneficial to first establish a framework for best practices. A helpful concept from the evaluation literature is the notion of the "core components" of a program. Core components are described as the essential principles of a program and their associated activities that are required to produce the changes intended by the program, as opposed to those features that can be modified without jeopardizing outcomes (Blase & Fixsen, 2013). Anchored by theory-driven but empirically-derived principles, core components include the operationalizing of the program principles into the contextual and structural aspects of the program and its specific intervention activities. The research evidence dictating the population or settings to which ISSP is best-suited defines the contextual aspects of ISSP. Best practice guidelines stipulating the ideal duration and intensity of the program would characterize structural aspects of the program, whereas best practices pertaining to the types of activities participants should engage in define the specific intervention activities.

The degree to which research is able to inform the various levels of core components varies. Few evaluation studies provide information about specific program components such as dosage, which is the amount of intervention received (Dane & Schneider, 1998). Moreover, given that randomized control trials and other rigorous research designs demand a high level of fidelity, it is unlikely that these studies are able to provide information about the functional elements of programs due to low levels of variation (Blase & Fixsen, 2013). Despite the potential for gaps in the research, the notion of core components provides a useful way of thinking about the best practices and the various levels on which they may be operationally defined, as well as a framework to structure and organize concepts.

Sources of best practices. Lipsey and colleagues (2010) identify principles extracted from meta-analyses of effective programs as the preferred source of evidence-based best practices. The amount of evaluation research on juvenile justice programs is extensive and the literature includes sufficient studies to provide indications of how much program elements such as intensity or participant characteristics might impact the effect sizes of specific types of programs (e.g., Lipsey, 2009). Where the literature exists, this information would ideally come from meta-analyses of similar types of programs to the one for which the best practices are being developed (e.g., James et al., 2013).

The adult meta-analysis literature on rehabilitative justice programming may also be a valuable source of best practices, particularly in the form of a theoretical framework for intervention. The Risk-Needs-Responsivity model (RNR; Andrews, Bonta, & Hoge, 1990) is well-established in adult rehabilitative programming and shows some support for youth (Andrews & Bonta, 2010b; Andrews, Zinger, Hoge, Bonta, Gendreau, & Cullen, 1990; Hoge, 2002). Generally speaking, the use of best practices suggested by the adult literature should be limited to those that are developmentally-appropriate for youth and should not be extended downwards to youth without caution. Risk factors for violence show changes even from late childhood to adolescence (Herrenkohl, Maguin, Hill, Hawkins, Abbott, & Catalano, 2000), such that treatment programs for adults may overlook or place less emphasis on risk factors that are important for youth (e.g., peer influence, school problems). Furthermore, inconsistencies have been found for effect sizes for RNR treatment principles between youth and adults. In the Andrews and Dowden (2006) meta-analysis, risk level did not impact treatment effect sizes for adults, but was significantly related to effect size for youth. Thus, prior to ruling out any best practices because they lack support in the adult literature, it is important to ensure that the adult findings are consistent with available evidence in the youth literature.

While findings from meta-analyses provide a highly useful starting point for developing or evaluating programs, they are necessarily broad and are not able to provide information on effective program components beyond general practices. Meta-analysis is limited in that it provides an average effect size of a program component or feature, for which there may be significant heterogeneity or variability in the effect size. As such, while a program component may be generally associated with positive results, it is possible that in certain programs its effect was negligible. Alternately, a program

feature that is associated with null results in a meta-analysis may be important to an individual program's success due to the presence of other program features not captured by the meta-analysis or many other reasons that contribute to a failure to yield a significant average effect. The ability of meta-analysis to provide best practices is moreover dependent on the information available in the research literature. Few studies provide good data about the use or quality of program components in practice, such that researchers are forced to use crude indicators or proxy variables to estimate the quality of the program's delivery for meta-analysis (Lipsey, 2009).

In all, meta-analysis is able to provide good data on a few program principles but the quality of the research literature at this time does not allow for further elucidation of effective components for many programs. This leaves juvenile justice agencies to fill in the gaps involved in the day to day operation of the program, for which less rigorous sources of evidence might be employed. Also, juvenile justice agencies may want to incorporate new or innovative practices in their programs while still recognizing that these should be supported by research and be evaluated. Non-quantitative systematic reviews of individual evaluation studies may be one source of evidence-based practices. Furthermore, in order to operationalize best practices, other forms of evidence may be required including findings from individual studies. Some of these promising practices do not have rigorous empirical support and are program elements that are commonly included in similar programs to Youth Justice ISSP as core components or that were identified retrospectively as elements that were lacking from unsuccessful programs. Nonetheless, these promising practices provide a starting point for empirical study.

Facets of Evaluation

Though best practices provide the foundation for an effective program, the program's delivery in practice may fall far short of the program design ideal. To the extent that programs diverge from evidence-based practice, both at the program design level as well as in the individual program delivery, the program is less likely to be effective. There are thus two primary aspects of program evaluation research: Implementation or process evaluation, which assesses whether the program is delivered as it was intended, and impact or outcome evaluation, which ascertains whether the program resulted in its intended aims.

Implementation. Along with studies of intervention outcomes, implementation is being recognized as a separate and important science. Unfortunately, research on the implementation of programs has lagged behind outcome research and there is a gap between knowledge of effective treatment or treatment principles and the actual programs that are received by consumers (Fixsen, Naaom, Blase, Friedman, & Wallace, 2005). For example, in a review of 158 experimental outcome studies of Applied Behavior Analysis, Gresham, Gansle, and Noell (1993) found that only 15.8% of studies reported systematically measured data on the implementation of the independent variables.

A lack of careful attention to the implementation of a program can mitigate the benefits of programs that have demonstrated effectiveness. As noted by the Office of Juvenile Justice Prevention (2004): "A poorly implemented program can lead to failure as easily as a poorly designed one" (p. 1). This assertion is supported by the finding in meta-analyses that better implemented programs yielded the largest effect sizes (Gresham et al., 1993; James et al., 2013; Lipsey, 2009; Wilson & Lipsey, 2005). Also, where a program is found to be effective but implementation quality is poor or inconsistent, the likelihood that outcomes can be attributed to the program is reduced (Mihalic, Fagan, et al., 2004).

Poor implementation poses an even greater threat to programs like ISSP that have yet to establish their effectiveness. A good program may be discarded in the face of outcome studies yielding null results for reasons other than a flawed program design. In a field study of two reintegration-oriented programs for adults with previous justice system involvement, only one in 20 participants received all four of the intended program components, while 41% of the "treatment" group did not receive any of the program's components (Dobson & Cook, 1982). In this case, it is clear that the outcome data alone would not have provided accurate conclusions about the program's effect.

Outcomes. Outcome evaluation is used to determine whether programs yield a change in the anticipated direction for the program participants (Mihalic, Fagan, et. al, 2004). While recidivism is generally the outcome of interest for juvenile justice programs like Youth Justice ISSP, collecting information on a broad range of outcomes is beneficial for a number of reasons. As ISSP-type programs target a broad range of needs, it is important to collect data on these outcomes in addition to recidivism to

determine if they are being effectively addressed through the program (Wiebush et al., 2005). Criminogenic needs such as substance use or poor school engagement serve as intermediate targets of change in the reduction of recidivism and are the mechanisms of criminal behaviour that rehabilitative programs target (Andrews & Bonta, 2010a). As such, it is important to measure changes in criminogenic needs since they will theoretically precede changes in reoffending, which may occur in the longer term outside of the follow-up period. Changes in outcomes such as school, employment, and community program participation are also important to measure as these experiences continue to influence youth after the program has ended (Altschuler & Armstrong, 2004).

The Current Study

It is not known whether Youth Justice ISSP is being delivered as intended or whether the program is contributing to positive change in participants. To yield the best possible outcomes, it is important that the program design is consistent with evidence-based practices and that the program delivery is of high quality. The primary purpose of this study was thus to investigate the implementation of Youth Justice ISSP as well as how the quality of implementation affected outcomes. A second purpose of the study was to develop and test an evidence-based framework of evaluation specific to ISSP that can be applied to individual-level data, drawing on similar models for general youth justice programming (Andrews & Bonta, 2007; Lipsey et al., 2010). Despite the evidence base for many of the specific best practices, it is important to evaluate these practices locally to ensure that they remain applicable under the circumstances operating for Youth Justice ISSP. It is also hoped that this study will provide much-needed data on the actual implementation of specific best practices, since most meta-analyses have had to rely on program averages or intended implementation rather than data from program delivery. Finally, findings for implementation could provide indirect support for the effectiveness of ISSP in yielding positive recidivism and other outcomes.

Proposed best practices for Youth Justice ISSP. The following is a summary of practices derived from meta-analysis and evaluation research that define an effective ISSP. The best practices that are the focus of the current study are grouped and identified according to the core concepts framework. Chapter 3 concerns the selection of youth for ISSP, Chapter 4 pertains to referral practices, Chapter 5 relates to

contextual aspects of the program, Chapter 6 examines structural features, Chapter 7 speaks to the program activities, while Chapter 8 provides an overall survey of the program's implementation according to program guidelines and best practices[1]:

- Select high-risk youth for the program[a]
- Choose an effective program type
- Interventions are guided by a therapeutic philosophy
- Programs are sponsored by the youth justice system
- Youth justice personnel are responsible for the delivery of programs
- Participation is mandatory
- Use a systematic risk assessment tool to identify youth[b]
- Identify goals for youth using a systematic risk assessment tool[b]
- Use individualized case planning
- Facilitate youths' transition to the community from custody
- Target individual and systemic factors using an individual modality[c]
- Make contact with ongoing or new community resources[c]
- Involve youths' families[c]
- Prioritize youths' transition to school or the workforce[c]
- The program is of sufficient intensity to yield an effect[d]
- Build long-term relationships[d]
- Contact occurs in periodic sessions
- Activities target criminogenic needs[e]
- Target criminogenic needs that are specific to the youth[e]
- Balance the support and supervision functions of ISSP[e]
- Use rewards and sanctions
- Implement the program properly according to the program design[f]

[1] Best practices marked with an a) are discussed in Chapter 3, b) in Chapter 4, c) in Chapter 5, d) in Chapter 6, e) in Chapter 7, and f) in Chapter 8

The first step in evaluating ISSP is to perform a cursory comparison of the program's guidelines to best practices to determine the correspondence to evidence-based practice at the program design level. The best practices reviewed briefly in the following section are practices that are consistent for all participants in the program or for which a file review methodology could not provide the data necessary to evaluate implementation. The best practices featured in Chapters 3 to 8 include both a program-level review and an evaluation of the actual implementation of the best practice among individual ISSP participants.

Reviewing Youth Justice ISSP's program design. Youth Justice ISSP includes comparable elements to other types of effective programs such as intensive probation programs, mentoring, and multiple services coordination (*choose an effective program type*; Lipsey, 2009). ISSP furthermore expressly notes the importance of incorporating a rehabilitative focus rather than exclusively serving the functions of punishment, surveillance, and public protection (*interventions are guided by a therapeutic philosophy*; Bonta & Andrews, 2007; Lipsey, 2009). Youth Justice is a branch of the Ministry of Children and Family Development, which funds and establishes the guidelines for Youth Justice ISSP (*programs are sponsored by the youth justice system*; Lipsey, 1999). The Youth Justice ISSP is delivered by custody staff (*youth justice personnel are responsible for the delivery of programs*; Lipsey, 1999). In order to ensure that Youth Justice ISSP is in place prior to release for participants in custody, the youth custody centre case management coordinator is able to make referrals in addition to community youth probation officers (*facilitate youths' transition to the community from custody*; Gies, 2003). The first duty listed for ISSP is that a case plan of support and supervision activities is developed in conjunction with youths' supervising probation officer (*use individualized case planning*; Gies, 2003).

Including the best practices reviewed in Chapters 3 to 8, the Youth Justice ISSP guidelines reflect approximately 70% of the best practices for similar types of programs. While no metric currently exists to gauge the extent to which a particular set of best practices might be expected to reduce recidivism, the high number of best practices suggests that the program's design has the potential to be effective based on the research evidence.

Evaluating Youth Justice ISSP's program delivery. There are two sets of standards against which the delivery of Youth Justice ISSP may be evaluated using program data. The first of these is the degree to which actual program delivery meets the relevant Youth Justice ISSP guidelines. The second standard concerns consistency with the best practice framework for ISSP-type programs. As with how these terms have been used elsewhere (e.g., Forgach, Patterson, & DeGarmo, 2005), *adherence* is used to refer to the program's delivery according to Youth Justice ISSP program guidelines and *fidelity* is used to refer delivery according to best practices.

The adherence and fidelity standards overlap, with the best practices subsuming most of the program guidelines. The focus of the current study for outcomes is on the best practice model for the program, given that the comprehensive model has the greatest likelihood of yielding positive outcomes. Although Youth Justice ISSP workers are not explicitly directed to implement the full best practice model by the ISSP guidelines, it is likely that some staff nevertheless applied these practices due to other prior knowledge of best practices, training, through their experience with the population, or for other reasons. For example, despite the absence of specific program guidelines indicating that participation in the program is mandatory, anecdotal file information in the current study suggested that youth were breached for not participating in ISSP (*participation is mandatory*; Lipsey, 1999). Furthermore, file data indicated that Youth Justice ISSP workers used rewards to acknowledge youths' graduation from programs or high school (*use rewards and sanctions*; Gies, 2003*)*. Unfortunately, information pertaining to these best practices was not recorded with sufficient consistency or specificity in the probation files to be reliably coded in the current study.

Hypotheses pertaining to individual best practices are detailed in their specific Chapters 3 to 8, which are grouped conceptually according to the core components as indicated in the footnote above. However, the primary hypotheses of the study are as follows:

Implementation hypotheses*.*

1. The actual delivery of the program will adhere to the Youth Justice ISSP program guidelines.

2. The best practices that overlap with the Youth Justice ISSP guidelines are expected to show high fidelity, but the best practice model is anticipated to show lower fidelity than the level of adherence to the Youth Justice ISSP guidelines overall.

Outcome hypotheses.

1. Higher adherence to best practices will lead to improvement in intermediate outcomes (e.g., school or vocational program enrollment, fewer instances of detected substance use, fewer breaches of conditions, increased program participation, etc.).
2. Higher adherence to best practices will result in reductions in recidivism outcomes in the year following ISSP completion (e.g., any re-offence, number of offence dates, types of offences, time in the community, etc.)

Chapter 2.

Methods

The current study used a multi-method approach focused primarily on a file review methodology. Retrospective methods were selected over a prospective evaluation in order to provide implementation data prior to a full evaluation. Given the substantial resources required for a prospective evaluation and the lack of data on the program's implementation, this study was a first step in determining the quality of the program's delivery and identifying potential areas of improvement prior to a prospective evaluation. A resource-intensive evaluation of a poorly-implemented program provides limited information regarding the program's true effectiveness (Goodstein & Sontheimer, 1997). Thus it was important that these issues be explored so that a prospective evaluation would capture the full potential of ISSP.

In the course of the study, several issues were encountered with the files that required decisions and additional procedures that are detailed in Appendix B (Appendix A comprises the tables and figures for the study). Also, a youth probation officer (YPO) survey was used to supplement gaps in the information gathered in the primary study. A Youth Justice ISSP staff survey was not completed due to the relatively small sample size anticipated and the difficulty of tracking staff, as ISSP positions rotate every two to three years. As such, few ISSP workers from the research time period of interest would still be in the position and the surveys would be biased towards more current cohorts.

Procedures

File review. Probation files were reviewed on-site at Metro Vancouver probation offices or sent via secure inter-office mail when located at sites outside of the region. Unfortunately, open files outside of Metro Vancouver could not be reviewed in the current study as they could not be shipped off-site. Demographic and background

variables, Youth Justice ISSP information, and non-recidivism outcomes were coded from probation files. Recidivism outcomes were coded from the Corrections Network (CORNET). Multiple recidivism outcomes were coded including number of convictions, offence types, and time in the community.

Data were extracted from the files using a coding sheet designed for this study. Two cases were used for pilot coding to determine what important variables according to the research literature could be extracted from the files. Recidivism data were generally recorded several months after the file coding or by a separate individual using a different coding sheet to maximize the likelihood that raters would be unaware of recidivism outcomes while collecting file and risk coding data.

Data coding. Best practices were coded from the extracted data. A coding manual was developed to rate the 13 program guidelines and best practices (available from the author upon request). For each best practice, a dichotomous variable was created from the research literature. The individual variable ratings were then added to derive a best practice composite. A 5-point variable was also created for each of the individual best practices to study their specific effects. The research literature provided the foundation for the variable definitions and the definitions were further shaped by the natural variation in the program's delivery to yield an approximately normal distribution. Further details on how the variables were operationalized are included in Chapters 3 to 8 and the coding manual, while the distributions of the variables are available upon request from the author.

Probation survey. Youth probation officers (YPOs) were recruited to participate in a survey study via an email request from regional Youth Justice Consultants to Probation Team Leaders that was then distributed to staff. Data collection involved online self-report questionnaires with items pertaining to demographic information, YPOs' role in ISSP, and program information. The measure was developed specifically for this study and comprised both closed- and open-ended questions (survey available from the author upon request). By way of compensation, YPOs were mailed a $20 gift card to a restaurant, coffee shop, or grocery store of their choice.

Participants

File review sample. Youth were selected for the study using a random number generator from a spreadsheet of all youth participating in ISSP generated by British Columbia Youth Justice. The Youth Justice probation files of 176 male and female youth under the age of 18 at the time of program start and who had completed ISSP between July 31, 2008 and September 30, 2011 were reviewed to allow a minimum one year follow-up period for outcomes. A year-long follow-up period subsequent to completion of the program was recommended by Altschuler and Armstrong (1994).

Youth who participated in Youth Justice ISSP from all three custody centre regions were included (Burnaby, Victoria, and Prince George) to provide a representative sample of the youth receiving the program in British Columbia, although region was coded as a variable to account for systematic differences between these areas. Youth receiving ISSP from contracted community providers were not included in the study due to the logistic constraints of securing permission from these agencies as well as diversity in program values, resources, and other factors that would create greater variability and further complicate the study of ISSP's implementation.

The sample was selected using a stratified random sampling procedure to ensure adequate coverage from each custody site. The final proportions did not precisely match those of the overall Youth Justice ISSP population due to some files being excluded subsequent to sample selection and these exclusions disproportionately affecting the Burnaby site. In particular, files that were excluded due to youth receiving duplicate community ISSP services did not appear to be a problem at the other sites. However, the final proportions of the sample were roughly equivalent to the original population.

Estimated sample size required. The estimated sample size required for this study was calculated based on the hierarchical regression and logistic regression analyses, as these were the primary analyses used in the study. The principal research question was whether or not implementation would be associated with recidivism and other outcomes after controlling for other factors of interest (e.g., risk, demographic variables). An effect size of $f^2 = 0.033$ was initially estimated for power analyses for implementation based on Lipsey's (2009) meta-analysis findings for multiple services programs, which is considered a small effect size (Cohen, 1988). With an estimated 5

A-set (control variables) and 1 B-set (implementation) predictors in the model, the sample size estimated to provide adequate power to detect an effect was 239 (Soper, n.d.). For logistic regression, the formula of N = 10 k/p was used, where k = number of predictors (6) and p = proportion of negative cases for any offending (0.58) during follow up by the eligible sample was used (Peduzzi, Concato, Kemper, Holford, & Feinstein, 1996). The formula yielded an estimated sample size of 104 participants.

The higher of the required sample sizes was revised in the current study for several reasons. The effect size itself was likely to underestimate the effect of implementation due the use of proxy variables to represent the quality of service delivery in the meta-analysis (Lipsey, 2009). Given the relatively more precise measures of implementation in the current study, a larger effect size was anticipated. Also, meta-analysis yields a mean effect size, while the effect size for implementation may vary widely across studies. Although Lipsey (2009) did not report the degree of heterogeneity in the effect sizes for implementation, the James and colleagues' (2013) meta-analysis found considerable heterogeneity in the effect sizes for implementation of aftercare programs (Q = 64.87, df = 3, p < 0.001). With the wide variety of formats associated with multiple services programs, it might be anticipated that a program that is relatively unstructured like Youth Justice ISSP would yield higher effect sizes for implementation. Furthermore, while still meaningful, a small effect size may not have sufficient clinical utility to justify the costs required to implement best practices, including training costs and time away from youth.

Thus, assuming an effect size of 0.05-0.10, which is still in the small-medium range by Cohen's (1988) standards, and the other parameters as noted above, the study was estimated to require between 80 and 160 participants to detect an effect. Also, a sample size in this range was consistent with the sample size required for the logistic regression analyses, which were used for most of the individual best practice analyses. Nevertheless, participants were oversampled due to issues with the data quality, which are described in more detail below.

Excluded participants. British Columbia Youth Justice supplied a data file of 536 youth who participated in Youth Justice ISSP during the study period. A significant proportion of these youth were not included in the study due to ineligibility and other factors detailed in Appendix C. Group differences in the characteristics of participants

are presented in Tables 1 and 2 and are discussed in Appendix D (as missing data). Briefly, there were differences among some of the groups that were anticipated based on the stratified sampling procedure and exclusion criteria that disproportionately affected sites, but this did not appear likely to impact the study findings. Participants who were excluded from the study due to legal disclosure restrictions after a period of non-offending were, not surprisingly, significantly less likely to have offended after ISSP start across multiple measures of recidivism (see Table 2). It was unfortunate that these youth were excluded since their successful completion of the non-disclosure period may have been associated with their ISSP participation. Their exclusion may thus have attenuated the effect size for ISSP implementation. These youth were otherwise similar to the included youth on demographic and offence characteristics prior to ISSP.

Final sample characteristics. The mean age of youth in the study was 16.23 years. The sample of youth was 77% male and 23% female. The ethnic composition of the sample according to the database was approximately 65% White, 10% Aboriginal, 4% Asian, 3.5% Black, 2% South Asian, and 13.5% other. Approximately 23% of the sample was identified as having some Aboriginal background, while 44% of the sample was identified as "unknown" with respect to Aboriginal identification according to the database provided by Youth Justice. The sample of youth was drawn from the Burnaby (61%), Victoria (25%), and Prince George (14%) catchment regions.

Study youth were convicted of an average of 6.50 (SD=5.93) and a median of 4 offences prior to their participation in Youth Justice ISSP, while the actual number of convictions ranged from 1 to 27. Thirty-eight percent of youth did not have any convictions prior to the offence(s) that led to their referral to ISSP. The majority (81.2%) of youth had been convicted of at least one violent offence (including sexual offences) prior to commencing Youth Justice ISSP, while 59.1% of youth had been convicted of at least one indictable offence, and 61.9% of youth had been convicted of at least one breach-related offence.

During the year follow up period after completing ISSP, 42% of participants were convicted of an offence. Youth were convicted of an average of 1.41 (SD=2.33) and a median of 0 offences during the year of follow up, with the number of convictions ranging from 0 to 12. Only 21% of youth were convicted of one or more violent offences, 12.5%

were convicted of at least one indictable offence, and 27.3% were convicted of one or more breach-related offences.

See Tables 3 for the characteristics of the study sample of Youth Justice ISSP participants. Note that the differences in mean age at the start of ISSP and offences for the sample from the information reported in Tables 1 and 2 are due to discrepancies between database and file information.

Probation survey. A total of 36 YPOs who had supervised youth who participated in Youth Justice ISSP in the previous five years completed the survey. Requests for the total number of YPOs at sites that were sent by email were not returned, such that the response rate is not available.

Respondent characteristics. Survey respondents were 38.9% male and 61.1% female. Participants reported an average of 13.04 (*SD* = 5.23) years of experience as a YPO, with a range of 2 to 25 years. YPOs reported supervising an average of 22 (*SD* = 29) youth participating in Youth Justice ISSP, ranging from 2 to 150 youth.

Data Quality

As is often the case with relying on retrospective file data, there were a number of ways in which the available information was less than ideally suited to a research study. These problems fell under three major categories: Missing data, inadequate data, and inaccurate data. Of particular relevance to this study, an unforeseen problem was the inconsistent number of logs on file and the poor quality of detail in the program summaries for the purposes of the study. A complete discussion of the measures taken to mitigate the threats to validity from these problems is included in Appendix D.

Briefly, missing data were dealt with using listwise deletion as there were generally few missing (e.g., <3%) or the data appeared to be missing randomly (Roth, 1994). In the case of the risk instrument used for the analyses (Structured Assessment of Violence Risk in Youth; SAVRY) and best practice composite, values were missing for 7.4% and 20.5% of cases, respectively. Missing data analyses were conducted to test whether the likelihood that cases were missing was associated with the outcome or predictor variables used in the study. When the means for the outcome variables were compared with the SAVRY total and best practice composite missing cases removed

and included, respectively, the mean differences and *t*-values were small, supporting a random missing data mechanism. Where listwise deletion was used, it is identified and the final sample reported.

Inadequate data, for which entire data sources were missing, and inaccurate data were managed by using the best available information, as there were often multiple sources of data that occasionally conflicted. Of particular importance to the study, ISSP logs were missing entirely from files for approximately one-quarter of youth, while the average proportion of logs on file was for 40% of the months that youth had spent in the community during the program. Even fewer logs provided information about specific program activities, focusing instead on updates about participants' circumstances. A number of files contained probation notes detailing ISSP activities that were able to supplement log information. Nevertheless, although it had been anticipated that ISSP logs would be the primary source of information for coding, these often proved to be lacking for the purposes of the current study.

Although the measures employed were carefully selected to minimize threats to validity, the data were nonetheless limited in their ability to address the research questions for the current study. In particular, conservative strategies were used to ensure that activities or effects were not falsely attributed to the program or best practices. However, the result of a conservative approach is the likelihood that already small effect sizes for implementation quality (e.g., Lipsey, 2009) would be further attenuated.

Inter-rater Analyses

See Appendix E for a detailed explanation of the inter-rater procedures used in the current study. Ten files were selected for inter-rater analyses for the data extraction process as well as the file outcome variable coding and risk ratings. Inter-rater agreement analyses were run first to determine the level of agreement between the data extracted from the files by the author and the two coders in the study prior to determining whether there was sufficient reliability for the variables operationalized for the study. Raters met or exceeded the cut-off for excellent reliability for dichotomous and continuous variables. Had the inter-rater agreement been low for data extraction

generally speaking, it would have been unlikely that a high level of reliability would be yielded for the program and outcome variables coded for the study.

For the risk ratings, the intra-class correlation coefficients (ICCs) for the Youth Level of Service/Case Management Inventory 2.0 (YLS/CMI 2.0) and SAVRY total scores fell in the excellent range, while the SAVRY structured professional judgment ratings fell in the adequate range. As a result, the latter was not used to control for risk in further analyses, although it was used to define the selection of high-risk youth variable due to a lack of cut-off scores available for the SAVRY. The reliability estimates were somewhat lower than previous estimates found for these measures in file review studies (e.g., Catchpole & Gretton, 2003; Viljoen, Elkovitch, Scalora, & Ullman, 2009); however, this was likely due to the small sample of files used for the inter-rater analyses due to the significant amount of time required to code files.

Nearly all of the outcome variables had sufficient reliability (i.e., coefficient estimates >0.60 or "good" reliability) to be included in the study, with several exceeding the cut-off for excellent reliability (Cicchetti, 1994). Variables that could not be coded with sufficient reliability and thus did not provide a valid index of the outcome were instances of arrests and unofficial offences documented in the file. These variables were excluded from the study, particularly given that official data were available for offending. The one exception where the dichotomous variable was retained despite poor reliability was for whether or not youth held any part-time employment during ISSP or the follow-up period. Since employment was a key outcome, the part-time variable was included in the study as may be the case when it is important to have such information (Långström et al., 1999). However, it should be noted that low reliability increases statistical noise and the likelihood of failing to find an effect where one exists (Hallgren, 2012). See Table 4 for the inter-rater estimates for the complete set of outcome variables.

For the best practice variable coding, due to its centrality to the study, precision analyses were conducted a priori to determine the number of cases necessary to obtain the desired confidence interval width around planning estimates derived from test-retest analyses. The number of cases required to obtain a reliability coefficient with two raters for which the lower end of the confidence interval would exceed the cut-off for excellent reliability (0.75) was approximately 20 (Bonnett, 2002). The inter-rater estimates for the

individual best practice variables and best practice composite are reported in Chapters 3 to 8. See Table 5 for the complete list of reliability estimates and confidence intervals.

Measures

Two empirically-validated measures were coded to provide an index of risk prior to ISSP participation, in addition to the Youth Justice probation risk assessment tool (Youth Community Risk/Needs Assessment; YCRNA). Although the YCRNA includes several factors that commonly appear in validated risk assessment tools, no evaluation studies could be located for the instrument. A study examining the adult version of the tool (Community Risk/Needs Assessment; CRNA) with a youth sample found it to have poor predictive validity relative to empirically-validated risk tools and low inter-rater reliability (Jack, 2000). The YCRNA was still used as an index of risk in the current study because probation officers have the best information to form a risk estimate; however, the two following measures were also rated by the file coders.

The Youth Level of Service-Case Management Inventory 2.0. The Youth Level of Service-Case Management Inventory 2.0 (YLS/CMI 2.0; Hoge & Andrews, 2011) is a risk/needs assessment tool for general offending in youth. It includes scales for offence history, education, substance abuse, family, personality and behaviour, peers, leisure and recreation, and attitudes and orientation. Each item is scored as absent or present, yielding a total score for risk and a risk level for each scale. Separate male and female norms are available to classify youth by risk level. There is furthermore a section for responsivity factors, although these were not included in the current study. The YLS/CMI 2.0 includes the same risk factors as the original version (YLS/CMI), on which much of the validity and reliability research has been conducted, but the YLS/CMI 2.0 has expanded the age norms to encompass youth from 12 to 18 years old and includes gender-informed and culturally-informed responsivity factors. The YLS/CMI was normed on a sample of 263 Canadian adolescent offenders, while the YLS/CMI 2.0 expanded the norms to over 12,000 American youth (Hoge & Andrews, 2011).

While the YLS/CMI 2.0 is relatively new, a substantial number of independent studies have found the YLS/CMI to have good reliability and predictive validity for general and violent offending for a range of populations including males and females, youth from diverse ethnic backgrounds, and youth who have committed sexual offences

(Jung & Rawana, 1999; Olver, Stockdale, & Wong, 2012; Onifade et al., 2008; Viljoen et al., 2009; although see Schmidt, Hoge, & Gomes, 2005; and Bechtel, Lowenkamp, & Latessa, 2007, for differential validity for female offenders). Nevertheless, evidence-based practice indicates that tools should be validated in the populations in which they are used. Fortunately, research studies are available that have been conducted with youth at the same sites as those in the study and confirm that the YLS/CMI shows strong predictive validity in this population (Catchpole & Gretton, 2003; Jack, 2000).

The Structured Assessment of Violence Risk in Youth. The Structured Assessment of Violence Risk in Youth (SAVRY; Borum, Bartel, & Forth, 2006) is a structured professional judgment instrument. The SAVRY includes 24 items in the areas of historical, individual, and social/contextual risk. The items were selected based on their empirical association with violence and offending. Clinicians make a rating based on their assessment of risk informed by these factors, rather than a total score, although the total score is frequently used in research.

Like the YLS/CMI, the SAVRY has been subjected to considerable empirical scrutiny. Despite its focus on violence risk, the SAVRY has been shown to significantly predict both violent and general recidivism (Catchpole & Gretton, 2003; Welsh, Schmidt, McKinnon, Chattha, & Meyers, 2008; Schmidt, Campbell, & Houlding, 2011; Viljoen, Scalora, Ullman, Cuadra, Bader, Chavez, & Lawrence, 2008; Vincent, Chapman, & Cook, 2011). A number of studies have found the inter-rater reliability for both the professional judgment and total score to be good to excellent (see Borum, Lodewijks, Bartel, & Forth, 2009 for a review).

Data Analysis Procedures

The primary analyses used hierarchical logistic regression for the 5-point individual best practice variables or best practice composite and the dichotomous variable for recidivism in the year following ISSP. One participant who was in custody for the entire follow-up period was excluded from the analyses. Demographic and psychosocial variables that were correlated with the recidivism variable were included in the first block to control for these significant youth characteristics (see Table 6). For the dichotomous recidivism variable, the covariates were a dummy-coded variable for youth identified as having no Aboriginal background ($r = 0.20$, $p < 0.05$), youth with mental

health issues (r = 0.19, p < 0.05), youth with low cognitive functioning (r = 0.18, p < 0.05), and the SAVRY total score (r = 0.28, p < 0.01). The SAVRY was selected to control for risk because Youth Justice is currently using this tool. Other outcomes were analysed using hierarchical linear regression for continuous outcomes and hierarchical logistic regression for binary outcomes. Covariates for the other outcomes are identified in the relevant sections. Where covariates were also correlated with the best practice variables, moderator analyses were used to clarify the nature of the relationship between the variables. Moderator analyses were not conducted for correlations between covariates due to the complexity that would be added to the model and because the relationships between covariates and recidivism were not central to the study.

Assumptions for hierarchical linear and logistic regression analyses were verified using the appropriate procedures for each (Cohen, Cohen, West, & Aiken, 2003). No significant violations of assumptions were identified.

Chapter 3.

Best Practices for the Selection of Youth

Prior to considering other aspects of the program's delivery, it is important to ensure that the appropriate youth are receiving the program. Participants selected for Youth Justice ISSP should be similar to the populations from which the best practices are drawn to increase the likelihood that the practices will be valid, given that characteristics of the population can moderate the effects of ISSP-type programs (James et al., 2013). This chapter pertains to best practices informing the type of youth who should be selected for the program.

The research literature indicates that higher intensity programs like ISSP should be targeted at high-risk youth (Andrews et al., 1990; James et al., 2013; Lipsey, 2009). Practically speaking, high-risk offenders are likely to make greater gains from programming because they have more criminogenic needs than low-risk offenders and it is thus possible to yield more significant improvement (Bonta & Andrews, 2007; Lipsey, 1995). Low-risk youth are likely to cease offending regardless of the programming that they receive, wasting the resources of an intensive program like ISSP, and they may occupy spaces in an oversubscribed program that prevents high-risk youth from accessing it (Andrews et al., 1990). In fact, due to the closer supervision provided by intensive re-entry or probation programs, offering such programs to low-risk youth may actually increase recidivism, as minor infractions such as curfew breaches are more likely to be detected (Altschuler & Armstrong, 1999).

In terms of the youth who are referred to Youth Justice ISSP, the program guidelines identify specific criteria for the selection of participants. Youth who commit first- or second-degree murder or a serious violent offence resulting in a custody sentence are the highest priority candidates for ISSP. Next are adolescents guilty of an offence of causing personal harm or attempted harm who are classified as high risk by

the YCRNA and who are released before the end of their custody term. Youth may also be sentenced to ISSP by a Court order (ISSO) under the third priority level. Where capacity allows, youth may be referred if they are given a conditional community sentence (e.g., Deferred Custody and Supervision Order), if they are on regular probation for a violent offence and deemed high risk, if they are released from custody early, or if they are on the community supervision portion of a custody sentence. The lowest priority youth are those on bail where circumstances warrant (e.g., high risk).

From the criteria listed above, it is clear that the ISSP guidelines for the selection of youth are based primarily on offence and sentence type. However, it is important that high risk is not conflated with the seriousness of the youth's offence, which is more likely to be reflected by the ISSP criteria. While past behaviour can be a predictor of future behaviour (Lipsey & Derzon, 1998), certain youth who commit serious offences actually have a lower recidivism rate than other types of offenders (Lattimore, Visher, & Linster, 1995). In particular, youth who have committed sexual offences tend to have fewer risk factors for general reoffending and lower recidivism rates than youth committing non-sexual offences (van der Put, van Vugt, Stams, Dekovic, & van der Laan, 2013), despite the fact that the offences are considered serious.

Another important issue for ISSP is how high-risk youth may be defined. Despite the availability of a local systematic risk assessment instrument (YCRNA) at the time youth were participating in ISSP, the use of a standardized tool such as the YLS/CMI 2.0 (Hoge & Andrews, 2011) or SAVRY (Borum, Bartel, & Forth, 2006) that has been validated by research conveys several advantages for ascertaining risk level. In that there are reliability and validity data available for these instruments, there is an empirical basis for the likely accuracy of the estimates (Hoge 2002). Although the YCRNA includes a number of risk factors supported by research, other tools developed using empirical research have failed to meet standards for correctly classifying youth when validation studies are conducted (e.g., Krysik & LeCroy, 2002). As with treatment, following empirical principles does not guarantee a positive outcome and it is necessary to evaluate a tool prior to relying upon it to provide an index of risk, which was not the focus of the current study. Furthermore, the use of an independent measure of risk in the study allows for a level of consistency and availability of estimates that may not be achieved when the YCRNA is used across probation officers in practice.

Evaluating program design and delivery. With respect to best practices for the selection of youth, Youth Justice ISSP was designed for high-risk youth, although the selection criteria are primarily prioritized according to the seriousness of the offence and are thus not entirely consistent with best practice. Risk ratings are, however, a consideration at some priority levels and are used to prioritize youth within a level where there are insufficient staff to provide service to all referred youth.

The research questions of interest in this chapter from the perspective of implementation were thus whether youth selected for the program were high-risk despite the absence of risk-based guidelines. In addition, implementation according to program guidelines was examined; that is, whether youth participating in ISSP met the Youth Justice criteria for selection of youth for the program. The relationship between the best practice and Youth Justice ISSP variables was also of interest, since although the risk principle is well supported in the research literature, a high correlation might suggest that the Youth Justice ISSP guidelines provide an alternate and sufficient index of risk. The probation survey was used to elucidate the decision processes involved in selecting youth for ISSP rather than inferring this information from the file alone.

In terms of outcomes, the risk principle of the RNR framework dictates that high-risk youth should benefit more from ISSP than low-risk youth in terms of recidivism and other outcomes (Andrews et al., 1990). It was also of interest whether youth meeting Youth Justice ISSP selection criteria benefitted more from the program than youth not meeting criteria. Furthermore, it was important to confirm that the validated risk tools (i.e., SAVRY and YLS/CMI 2.0) predicted offending in the current population, since they were used to define the risk criteria for the best practice variables.

Procedures

Due to the discrepancies between the best practice and Youth Justice ISSP selection criteria for youth, two separate variables were created to explore the relationships with outcomes for each. Separate dichotomous variables were created for use in the best practice and the Youth Justice ISSP implementation composites, respectively.

Best practice variables. The dichotomous variable for the best practice of selecting high-risk youth for ISSP was coded based on the presence or absence of a

high-risk rating for either of the YLS/CMI 2.0 rating categories or SAVRY professional judgment. The 5-point program variable was coded using combinations of the two risk ratings that were defined to create an approximately normal distribution. The variable ranged from a very high and high rating on the YLS/CMI 2.0 and SAVRY, respectively, to a low rating on both tools. The $ICC_{(2,1)}$ value for this variable was 0.95. Since the dichotomous variable was not only used to calculate the total score but also to distinguish between risk levels for the current section, Kappa was calculated for the variable and yielded excellent reliability ($\kappa = 1.00$, $p < 0.001$).

Youth Justice ISSP variables. The priority guidelines outlined in this chapter were used to define the Youth Justice ISSP variables. These variables were coded based on the selection criterion identified on the referral form, although many forms did not have a section to enter these ratings, the section of the form was not completed, or there was no referral on file. Where this information was not present, the criterion level under which youth fell could generally be estimated by the proximity of offence dates, sentences, and YCRNAs to the referral or program start date. For the dichotomous program variable, if the individual met any of the Youth Justice ISSP criteria, the variable was coded as present. The 5-point variable was coded using the program criteria priority levels. Due to the low number of youth falling at the two highest priority levels, these were grouped together. Since the criteria were already set, no adjustments were made in the operationalizing of this variable. Youth meeting multiple criteria were given the highest possible rating. The $ICC_{(2,1)}$ value for this variable was 0.87 and $\kappa = 0.41$ ($p < 0.05$) for the dichotomous variable. The latter fell at the cut-off for adequate reliability. Nevertheless, due to the importance of including the variable, it was retained, with the caution that poor reliability may contribute to greater error (Hallgren, 2012).

Results

Descriptive data. Risk scores on the YLS/CMI 2.0 ranged from 4 to 41, with a mean of 28 ($n = 166$). In terms of the categorical distinctions provided by the YLS/CMI 2.0, this score falls within the “high” risk category for both male and female youth. Total scores using the SAVRY ranged from 3 to 48, with a mean of 31 ($n = 163$).

Implementation. With respect to the best practice variable, 85.4% of youth participating in ISSP in the study were rated as high risk on at least one of the YLS/CMI

2.0 total score or SAVRY professional judgment. Youth were rated high risk (or very high risk) on both measures for 39.4% of cases. For the Youth Justice ISSP guideline variable, 81.1% of youth met at least one of the priority criteria. Specifically, 11.4% of youth met the highest criterion (i.e., personal harm, high risk, & custody sentence), 45.7% were referred following an ISSO sentence, and 20.0% were referred under the third criterion (e.g., deferred custody order, probation order with personal harm, etc.). Seven youth (4.0%) appeared to have been referred to ISSP while on bail, although the documentation was often particularly unclear for this criterion.

Probation survey data. Though the file study was able to provide information about the youth who participated in ISSP, it was necessary to infer conclusions about the decision-making processes leading to youth's referral and whether youth were being selected based on risk or ISSP criteria rather than other factors. However, descriptive data from the probation survey were able to add useful information about YPOs' decision-making.

Nearly all YPOs indicated that they considered risk ratings (YCRNA or SAVRY) to be either very important (50.0%) or somewhat important (36.1%) in their decision to refer youth to ISSP. YPOs appeared to rely less on the Youth Justice ISSP priority guidelines in determining appropriateness for ISSP, with 19.4% of respondents rating them as very important and 52.8% rating them as somewhat important, while 8.3% indicated that they were not at all important. However, YPOs did consider the type and seriousness of offences, on which the priority guidelines are largely based, to be very important (58.3% and 69.4%, respectively) or at least somewhat important (38.9% and 27.8%) in their decision to refer youth to ISSP.

Furthermore, 61.8% of respondents indicated that offence-related considerations (e.g., number, type, seriousness) were the most important factor in their decision to refer youth to ISSP. In contrast, 11.7% YPOs listed risk as the most important factor in the decision to refer to ISSP, with a further 20.6% ranking risk in their top three considerations. Only 11.8% of YPOs considered the fit with ISSP priority guidelines to be the most important factor in the decision to refer to ISSP, while 8.8% rated the guidelines to be in their top three considerations. Perhaps not surprisingly, 8.3% of YPOs indicated they were somewhat unfamiliar with the guidelines with a further 33.3% reporting being only somewhat familiar, although the majority of YPOs (55.6%)

considered themselves to be very familiar with the guidelines. The majority of YPOs felt that the Youth Justice ISSP priority guidelines "sometimes" (38.9%) or "often" (41.7%) allowed them to refer the most appropriate youth to the program, while fewer felt the guidelines did "not often" (2.8%) or "almost always" (16.7%) allowed for the most appropriate youth to be referred.

Best practice and guideline correspondence. In terms of the relationship between the risk tools and the Youth Justice ISSP selection criteria, the totals for each of the risk tools were significantly correlated with the 5-point variable for the program guidelines, suggesting some correspondence between risk and the mostly offence-based ISSP criteria. The correlation between the ISSP criteria and the YLS/CMI 2.0 was $r = 0.27$, $p < 0.01$, while for the SAVRY it was $r = 0.30$, $p < 0.001$. Although the correlations between the risk tools and ISSP guidelines were significant, for comparison purposes, the correlation between the YLS/CMI 2.0 and SAVRY was much higher, $r = 0.86$, $p < 0.001$.

Predictive validity. As a first step for the outcome analyses, the predictive validity of the SAVRY and YLS/CMI 2.0 risk tools was examined, since the risk tools provided the foundation for the best practice variable. Each of the YLS/CMI 2.0 total score (AUC = 0.81, $p < 0.001$, 95% CI 0.74 - 0.89), SAVRY total score (AUC = 0.75, $p < 0.001$, 95% CI 0.67 - 0.84), and SAVRY professional judgment rating (AUC = 0.73, $p < 0.001$, 95% CI 0.64 - 0.82) significantly predicted whether or not youth re-offended during ISSP and the follow-up period. Analyses were also conducted for the follow-up period only, since offences during follow up were used for the primary analyses of the study (for which risk was used as a covariate). The YLS/CMI 2.0 total score (AUC = 0.67, $p < 0.001$, 95% CI 0.59 - 0.76), SAVRY total score (AUC = 0.63, $p < 0.01$, 95% CI 0.54 - 0.72), and SAVRY professional judgment ratings (AUC = 0.67, $p < 0.001$, 95% CI 0.58 - 0.75) also significantly predicted whether or not youth re-offended in the year follow-up period.

Outcomes. The relationship between the two selection variables and recidivism was examined by separately analysing offending prior to ISSP and following ISSP using the dichotomous variables. Youth were separated according to the best practice high and low risk and according to those meeting Youth Justice ISSP criteria and not meeting criteria, respectively. The number of offences in the year before and the year after ISSP

were used to provide a relatively consistent measure of time for all youth. Youth who were in custody for the year prior or subsequent to ISSP and who thus had no opportunity to offend were excluded from the analyses (n = 4). A Repeated-Measures Analysis of Covariance (ANCOVA) was used to determine whether risk level or criteria match was associated with pre-ISSP to post-ISSP changes in offending. ISSP length was included as a covariate to provide an index of maturation, since youth participating in the program for a longer period of time would also be relatively older at the beginning of the follow-up period and may be less likely to offend due to developmental changes.

There was a significant interaction for the within-subjects pre- and post-ISSP number of offences and the risk best practice (F(1,164) = 4.58, p < 0.01; see Table 7). The marginal means are plotted in Figure 1. Post-hoc paired t-tests were used to further elucidate the relationship between risk level and offending. There was a significant decrease in offending for the year pre- to post-ISSP for both high- and low-risk youth (see Table 9). An independent samples t-test of the change scores indicated that the mean change from pre-ISSP to post-ISSP for the high-risk youth (M = 2.39, SD = 3.09) was significantly larger than for the low-risk youth (M = 0.71, SD = 1.30), t(76.58) = 4.52, p < 0.001, equal variances not assumed.

The interaction between the within-subjects pre- and post-ISSP offending and the criteria match approached significance (F(1,168) = 2.89, p = 0.09; see Table 8). The estimated marginal means for matched and non-matched youth are presented in Figure 2. Post-hoc paired t-tests indicate that there was a significant decrease in offending for ISSP participants in the year following ISSP regardless of whether or not they met criteria (see Table 9). An independent samples t-test of the change scores of the youth matched and not matched to the ISSP guidelines indicated that the mean change for the matched youth (M = 2.27, SD = 3.15) was marginally higher than that of the non-matched youth (M = 1.55, SD = 3.15), t(99.47) = 1.91, p = 0.06, equal variances not assumed.

Discussion

The majority of youth selected for ISSP were classified as high risk by ratings made from the file for the current study. This finding is notwithstanding that YPOs were unlikely to have had access to YLS/CMI 2.0 or SAVRY ratings since the SAVRY was

implemented subsequent to the study period, although the YCRNA has many factors that are the same or similar to those found on validated tools. As such, most youth who participated in ISSP were appropriately selected for the program according to the best practice for risk level on at least one of the tools. It was impossible with the information available to determine whether any high-risk youth were not referred or selected for ISSP or whether the few low-risk youth participating in ISSP might have occupied needed program spaces. The probation survey further confirmed that risk is an important consideration in YPOs' referral of youth to ISSP, although the majority did not identify risk as the most important consideration.

Most youth also met at least one of the Youth Justice ISSP priority criteria; however, nearly 20% of youth did not appear to meet criteria. This elevated number of youth may have been due to the quality of the data used for coding the variable or could reflect the referral of youth for other reasons that were believed to be legitimate. The findings were consistent with the relatively lower importance placed on ISSP guidelines reported by YPOs in the survey. Offence-related considerations, which form the basis of many of the criteria for the ISSP guidelines, appear to be the predominant factor driving referrals to ISSP according to YPOs.

Regarding the relationship between the best practice and ISSP guideline selection variables, although high risk and high priority were correlated, the correlation size was only in the small to moderate range according to Cohen's (1998) conventions for correlation effect sizes. The finding further underlines the fact that, while seriousness and type of offence may be considerations in determining risk, the constructs of risk and offence seriousness do not share much overlap. It is encouraging that in the absence of specific criteria regarding risk level, the program nevertheless appears to have been delivered to high-risk youth. Had the risk and ISSP priority variables provided a roughly equivalent measure of risk, it might have been beneficial to maintain the current guidelines as staff are already familiar with these guidelines and the use of these guidelines would not require additional training on the standardized risk tools (note: At the time of writing this document, Youth Justice had implemented the SAVRY for use by YPOs; P. Bartel, personal communication, September 2013). Also, the Youth Justice ISSP guidelines provide a quick means of determining appropriateness for the program, rather than requiring staff to complete a risk assessment prior to referring youth, as an

expedited referral to ISSP was identified by a number of YPOs (55%) as the highest priority. These considerations are important in determining whether to make changes to existing guidelines.

In terms of outcomes, both the YLS/CMI 2.0 and SAVRY total scores significantly predicted offending during ISSP and the year of follow up. According to Rice and Harris' (2005) guidelines, the AUC effect sizes associated with the SAVRY and YLS/CMI 2.0 for offending on ISSP and during the follow-up period indicate a high level of predictive validity. The effect sizes for the follow-up period alone were still significant but were somewhat attenuated relative to the time period that included ISSP, likely due to the program lasting a year or more for some youth and the dynamic nature of risk over time in adolescents (Herrenkohl et al, 2000; van der Put et al., 2012). It is recommended for both risk tools that assessments are repeated at regular intervals (Borum et al., 2006; Hoge & Andrews, 2011). Nevertheless, these findings support that the tools were able to identify high-risk youth in this sample and thus constituted a valid basis for the best practice variable used in the subsequent analyses.

Whether or not youth were high risk or met ISSP criteria, youth acquired fewer convictions in the year subsequent to their participation in ISSP than in the year prior. Unfortunately, without a control group, these reductions cannot be attributed to ISSP, although the results are more promising than if there had been no change on average or an increase in offending. Consistent with the risk principle (Andrews et al., 1990), significantly greater decreases in the number of convictions in the year following ISSP relative to the year prior to ISSP were observed for high-risk youth. Youth appropriately referred according to the ISSP selection criteria had marginally greater decreases in the number of convictions. The relatively low inter-rater reliability for this variable, likely due to the data quality, may have attenuated the relationship between criteria match and decreases in offending.

Some limitations of the current study for these research questions were that the risk tool raters were not blind to the status of the youth as participants in ISSP and may have been more liberal with higher risk ratings for this reason. Moreover, in an effort to maximize the number of valid risk assessments, item ratings were sometimes based on partial or limited information, which could lead to bias both by potentially lowering risk for youth with little information on file and by possibly elevating the risk of youth where only

a single example of an item may have appeared in the file. Also, without being able to selectively provide file information prior to ISSP to raters, there is the possibility that raters accidentally encountered information about offending subsequent to ISSP while rating items or by other indicators, such as the thickness of the file or number of volumes. Unfortunately, these problems were unavoidable in the current study due to logistical constraints.

The lack of a control group was particularly problematic for this best practice, given the difficulty of disentangling risk and other factors when determining whether high-risk youth benefitted more from programming. Given that high-risk youth are likely to have higher numbers of offences in the year prior to ISSP, they may show greater decreases in offending due to factors such as regression to the mean (i.e., extreme scores are likely to return to normal, while low scores are more likely to increase) or due to the simple fact that they have more potential for decrease than a low-risk youth who may have only committed one offence prior to their referral to ISSP. As a result, it is not clear whether high-risk youth benefitted more from the program or whether other factors contributed to the decreases in offending.

Recommendations. It may be beneficial for Youth Justice to revise the ISSP selection criteria to be consistent with best practice, prioritizing youth with a high risk rating using the SAVRY rather than criteria based primarily on offence or sentence characteristics. For example, a small number of youth with a sexual offence conviction were otherwise rated low risk and participated in ISSP for between 12 to 23 months. These youth may be more appropriately served through community agencies and a less intensive supervision program rather than consuming the significant resources provided by Youth Justice ISSP for the lengthy duration of their community sentences. YPOs themselves do not appear to feel strongly that the ISSP guidelines allow for the identification of the most appropriate youth, although some comments suggest that YPOs would not feel that criteria based on risk would be any more helpful. For example, one YPO indicated that, despite being familiar with research about ISSP having the potential to be counter-productive or intrusive for low-risk youth, it had been helpful for some low-risk youth to have short-term, focused participation in the program.

While it is unlikely that referring low-risk youth to ISSP would be as harmful as with other youth justice outcomes (e.g., custody or placement decisions; Lowenkamp &

Latessa, 2004), low-risk youth may still be occupying resources that are needed for high-risk youth. Low-risk youth should not be included in ISSP group outings with high-risk youth or enrolled in programs that will bring them into contact with other justice-involved youth (Andrews & Bonta, 2010a). Also, while ISSP workers may be tempted to assist low-risk youth directly with a number of their need areas, it may be helpful instead for ISSP workers to use their time to connect low-risk youth to community agencies that can provide assistance with participants' needs on an ongoing basis. Furthermore, Youth Justice ISSP should be treated as a last resort for low-risk youth, when non-justice and community options for a youth worker have been exhausted. Should Youth Justice deem the inclusion of youth not meeting criteria to be acceptable where caseloads allow, it would be important that youth are clearly identified as low risk on the referral and that YPOs coordinate with the ISSP worker to ensure that youth receive a short-term, less intensive program.

Referral criteria based on the risk best practice may also be useful to ease the referral process, since the current ISSP criteria are somewhat complex. In the current study, for example, only approximately 50% of YPOs identified the priority criterion under which they were referring youth on the referral. Regardless of whether the selection criteria for ISSP are changed, clear and consistent guidelines about the referral criteria should be conveyed to YPOs on a regular basis. As one survey respondent noted, "every once in a while it would be nice to be reminded of the guidelines."

Chapter 4.

Best Practices for the Use of Risk Assessment for Referral

While the best practices in the previous chapter define the "who" for Youth Justice ISSP, the best practices pertaining to the referral process concern "how" these youth are identified. As the first step of ISSP, how the referral process is navigated sets the foundation for the proper selection of youth and for which criminogenic needs are to be targeted. As the gatekeepers for youths' entry to Youth Justice ISSP, YPOs' referral practices are an important area of study. Although most Youth Justice ISSP criteria do not require that youth be identified as high risk to be referred (e.g., in the case of youth receiving an ISSO from the Court), best practices nonetheless indicate that risk assessment should be used guide the referral process.

Following from the previous chapter, it is important that youth are properly classified on the referral according to their risk level, since youth at high risk of reoffending benefit more from programs than low-risk youth (Andrews et al., 1990; Lipsey & Wilson, 1998). In the juvenile justice system, risk classification has tended to be done in an informal, unstructured manner, according to survey research on practices (Wiebush, Baird, Krisberg, & Onek, 1995). These classifications are generally drawn from such sources as the details of the case, probation officers' understanding of the causes of offending, and their experience with similar offenders (Van Voorhis & Brown, 1997). However, unstructured clinical judgment is plagued by problems with validity and reliability (Ziskin, 1995; Lidz, Mulvey, & Gardner, 1993; although see Mossman, 1994).

Rather than relying upon clinical judgment, suitable youth should be identified for probation programs as high risk using a systematic risk assessment instrument (Gies, 2003; Wiebush et al., 2005). Reasons for this include ensuring that YPOs do not include irrelevant or exclude pertinent information, that the relevance of information is weighed

equally among YPOs, and that YPOs' personal opinions will not shape the end result (Ballucci, 2012). For example, a study of adult probation officer practices found that participants tended to focus on the severity of offences and the length of offenders' criminal records in construing risk in their recommendations, while other risk factors such as substance abuse played a less prominent role (Giles & Mullineux, 2000). These findings are consistent with the key considerations for YPOs identified in Chapter 3.

Other problems with failing to use a systematic risk assessment instrument are that classifications rely on decision rules that may not be known or available to others and that practitioners may fail to consider the base rates of behaviour (Bonta, 1996; Gottfredson & Moriarty, 2006). A study examining YPOs' risk estimates before and after the implementation of a systematic risk instrument found that YPOs perceived a significantly lower percentage of youth to be likely to re-offend after being trained to use the risk tool (Vincent, Paiva-Salisbury, Cook, Guy, & Perrault, 2012). Moreover, the estimates of re-offence rates yielded from the risk tools were closer to actual rates of reoffending than those made by the YPOs prior to training on the risk tools (Perrault, Paiva-Salisbury, & Vincent, 2012). In a pre-post study design, the implementation of a risk tool led to better decision-making in other contexts, including the more appropriate use of out-of-home placements, supervision levels, and service allocation according to risk principles (Vincent, Guy, Gershenson, & McCabe, 2012). It is thus important to determine the extent to which YPOs use risk assessment in referring youth for ISSP, as research would suggest that the use of risk assessment could improve the identification of appropriate youth for the program.

A second function of the Youth Justice ISSP referral process is to provide some initial direction to ISSP workers about critical need areas for youth. YPOs again play an important role in informing ISSP workers about potential targets for the program, given that they often have a pre-existing relationship with youth and their families. The targets identified should adhere to the needs principle of the RNR framework to yield the largest reductions in recidivism (Andrews et al., 1990; Hoge, 2002). That is, when YPOs are identifying goals on the referral form, the specific criminogenic needs that are relevant to the adolescent's offending behaviour should be the focus of the program.

While the objective of targeting youths' needs appears to be straightforward, justice-involved youth often evidence a myriad of criminogenic and non-criminogenic

needs (Hoge, Andrews, & Leschied, 1996; Thompson & Pope, 2005). As such, YPOs and ISSP workers have a difficult role in correctly identifying and often prioritizing adolescents' criminogenic needs. A study of juvenile justice professionals in the United States indicated that only 36% of professionals identified at least one of the "Big Four" criminogenic needs (criminal history, antisocial attitudes, associates, and personality; Andrews, Bonta, & Wormith, 2006) thought to be most predictive of antisocial behaviour (Flores, Russell, Latessa, & Travis; 2005). The use of the probation risk tool, the YCRNA, should thus facilitate the identification and selection of appropriate criminogenic needs for ISSP.

Despite the availability of the YCRNA, a study of juvenile justice practices found that 86% of professionals used the YLS/CMI to inform risk decisions but only 56% relied upon it to identify treatment goals (Flores, Travis, & Latessa, 2004). Another study of case management practices found that, while YPOs generally classified need areas accurately according to risk assessment findings, they tended to over-classify certain need areas for intervention (e.g., making education recommendations when youth were low risk in these areas) and under-identify others (Luong & Wormith, 2011). The correspondence between needs and interventions identified was associated with significant reductions in re-offending, while the failure to identify needs for intervention was associated with increases in re-offending for high-risk offenders. As such, although there is a systematic risk tool in place for Youth Justice, it is important to study whether and how YPOs use the YCRNA to identify targets for ISSP.

Evaluating program design and delivery. As noted previously, since the referral program criteria are generally based on offence characteristics or sentences and most criteria are not dependent on risk assessment ratings, the best practice of using systematic risk assessment to identify youth is not considered to be present for Youth Justice ISSP. Nevertheless, the guidelines do make reference to avoiding "net widening" (p. 2, Ministry of Children and Family Development, 2006) and reserving the program for high-risk youth. There do not appear to be specific guidelines for ISSP pertaining to the identification of targets using risk assessment other than the general stipulation that the program should aim to target criminogenic needs.

The hypotheses for this chapter are that, despite the lack of program guidelines, YPOs are anticipated to use risk assessment in their referral since there is a systematic

risk tool in place that YPOs are expected to complete and some ISSP referral criteria require the risk level of youth to be high or medium. To this end, there is a section on most versions of the referral form where YCRNA ratings are to be entered. It is also anticipated that YPOs will use the YCRNA to inform their choice of needs that will be the focus of ISSP, consistent with Luong and Wormith's (2011) results, since YCRNAs are meant to be routinely completed in their case management (Hannah-Moffatt & Maurutto, 2003). However, some variability is anticipated in YPOs' use of risk assessment to identify goals as there do not appear to be specific instructions for program targets. The two risk assessment variables are anticipated to be correlated as YPOs who use the YCRNA to identify the youth's risk level as part of the referral process should be more likely to refer to the YCRNA to identify youths' needs. These individual best practices are not anticipated to be strongly related to recidivism outcomes since the referral process is relatively removed from youths' program experiences.

Procedures

For these best practice items, the use of risk assessment referred to the YCRNA since it was the risk classification tool used by YPOs during the period of study. Since a high risk rating is not required for referral, the focus of the following items was not the presence of a high risk rating but the correspondence of the values submitted in the referral process with those indicated by the YCRNA. Thus, the use of risk assessment was inferred if the YCRNA date preceded the referral date and the risk ratings were consistent on both forms. Nine cases were not coded because an earlier version of the referral form did not include risk ratings and 18 cases did not have a referral form on file.

The dichotomous variable for the identification of youth using systematic risk assessment was rated as present if a YCRNA had been completed prior to youths' referral to the program and the risk ratings were consistent between the YCRNA and the referral form. The 5-point variable was coded based on the correspondence between risk ratings as well as various timeframes for the completion of the rating form (i.e., day of, within a month prior, any time prior, and after referral) based on the natural variation of time frames in the sample. The assumption underlying this item was that risk assessment ratings completed closer to the referral date were more likely to be used to

complete the referral and to best reflect youth's current level of risk. The $ICC_{(2,1)}$ value for the variable was 0.88, which is considered excellent.

For the identification of goals using systematic risk assessment, the dichotomous variable was coded based on the presence of at least one goal identified that was consistent with a need rated as at least moderately problematic on the YCRNA. The 5-point variable was coded based on three features: The recorded targets' correspondence with need areas captured by the tool, the appropriate prioritization of targets according to the YPOs' ratings of level of need, and the comprehensiveness of targets in identifying the range of needs evidenced by the youth. The item ranged from no needs identified on the form to at least 75% of youths' problematic needs identified, with no non-criminogenic needs included. For this item, since the timeframe of completion was captured by the previous item, a less literal understanding of the use of risk assessment was adopted and the item focused generally on whether needs were selected from the YCRNA and were specific to the youth based on the need ratings on the closest available YCRNA. The $ICC_{(2,1)}$ value for this variable was 0.42, which is considered adequate. Although this variable fell well below the desired cut-off of 0.75, it was included due to its importance to the study as a best practice.

Results

Descriptive data. Completed referrals were on file for 155 youth. YPOs identified a mean of 4.21 (*SD* = 1.78) different needs or goals on the referral form. The frequencies of needs identified as problematic for youth, the needs identified as targets on the referral, and the need areas appropriately identified, under-identified, and over-identified are presented in Table 10. Criminogenic needs from the YCRNA that were most often appropriately identified on the referrals (i.e., rated as moderately or very problematic and identified on the referral or rated as areas of strength or non-problematic and not identified on the referral) were substance use (60.0% of youth) and education/employment (54.8% of youth). The need areas most frequently under-identified (i.e., not identified as a target where youth were rated as having problematic functioning) were family relationships (69.5% of those with an identified need) and attitudes (75.6% of those with an identified need). Needs were also over-identified, that is, they were identified as goals on the referral when they were rated as not problematic

or areas of strength on the YCRNA. This was particularly true for educational and recreational goals (9.7% and 9.0% of youth, respectively). Forty per cent of referrals included at least one goal pertaining to non-criminogenic needs or supervision and compliance with probation conditions, while 37% included at least one target that was vaguely worded and did not provide useful direction to ISSP workers as to potential program goals (e.g., “reduce criminal behaviour”).

In terms of the match between needs and targets, referrals addressed a mean of 39.32% (*SD* = 24.20) of youths’ needs identified as moderately or very problematic on the YCRNA. Of the goals or targets identified on referrals, an average of 64.92% (*SD* = 31.80) pertained to criminogenic needs rated as problematic.

Implementation. Regarding the use of risk assessment to identify youth, 59.9% of the 147 eligible cases were rated as having the best practice present; that is, the risk ratings identified on the referral were consistent with the nearest YCRNA completed prior to the referral date. Another 19.0% of cases had at least one YCRNA on file prior to referral to ISSP but the ratings of the most recent YCRNA were inconsistent with the referral. The remaining 21.1% cases did not have a YCRNA completed prior to referral. In 72.7% of cases where ratings were inaccurate, risk level was overestimated (e.g., high was entered on the referral form when YCRNA risk was medium).

Regarding the use of risk assessment to identify targets, 92.3% of the 159 eligible cases had at least one goal pertaining to a problematic criminogenic need. For the 5-point variable, 25.8% of cases identified at least 55% of youths’ problematic criminogenic needs and did not include any non-criminogenic targets, while 28.4% identified fewer than 25% of youths’ needs as targets or identified more non-criminogenic or non-problematic need targets than problematic need areas.

Correspondence between referral variables. The 5-point variables for identifying youth and identifying targets using systematic risk assessment were correlated to determine whether YPOs who appeared to have used the YCRNA to make the referral also appeared to have used the YCRNA to select program targets. The two variables were not significantly related (n = 143; r = -0.02, *ns*).

Probation survey data. As with the best practices for the selection of youth for ISSP, the probation survey was able to provide additional information about the use of

risk assessment as it relates to the referral for ISSP. Since it was necessary to infer the use of risk assessment based on limited and incomplete data, the probation survey responses were particularly informative for this best practice.

Relatively few YPOs (13.9%) reported relying substantially on risk assessment (i.e., YCRNA or SAVRY) to identify clients for Youth Justice ISSP. The majority of YPOs reported not relying on risk assessment at all (27.8%) or minimally (25.0%). In contrast, only 8.3% of YPOs indicated that they did not rely on risk assessment at all in identifying targets for ISSP. Nevertheless, only 16.7% of YPOs indicated relying substantially on risk assessment to identify targets, while the majority reported relying on risk assessment minimally (44.4%). In terms of the importance of a recent risk assessment for ISSP referral (i.e., completed in the month prior to referral), 25.0% of YPOs considered a recent YCRNA to be very important, while the majority considered it to be not at all important (19.4%) or somewhat unimportant (22.2%).

Outcomes. The relationship between referral practices and recidivism was examined using hierarchical logistic regression. None of the covariates for recidivism identified in Chapter 2 were correlated with the identification of youth or targets using risk assessment variables, so moderator analyses were not necessary. Neither the use of the YCRNA to inform the referral decision or to inform the program goals was significantly associated with recidivism (see Tables 11 & 12).

Discussion

In the current study, the majority of YPOs appeared to use risk ratings in referring youth to ISSP and still more had completed at least one risk assessment prior to referral. It is possible that ratings were inconsistent between the YCRNA and the referral due to simple error or haste in completing the form. Furthermore, nearly all YPOs identified at least one criminogenic need target on the referral that was rated moderately or very problematic for the youth. As such, many YPOs appeared to have followed best practices in their selection of youth and of targets for ISSP in the absence of specific ISSP guidelines. Nevertheless, the YPO survey responses would suggest that many YPOs do not consider risk assessment to be important to the referral process.

Surprisingly, whether YPOs used risk assessment to identify youth for the program was not related to whether YPOs used risk assessment to select program

goals. This finding was contrary to expectations that YPOs who had more recently completed the risk tool and appeared to use it when making the referral would be more likely to refer to the risk tool or have the information more readily available in memory to complete the targets on the referral form. However, the lack of correspondence is consistent with findings from Flores and colleagues (2004) that probation officers' decisions about treatment goals often did not appear to be informed by risk information.

Another observation from the current study was that in cases where there were inconsistencies between the YCRNA and the referral, most had an elevated risk rating relative to the YCRNA. It may be that YPOs elevate risk ratings in hopes of increasing the priority level for youth as other research has found overrides to commonly occur in structured decision-making among juvenile probation officers due to disagreements with the recommendation or practical constraints (Shook & Sarri, 2007). Another explanation is that YPOs may actually overestimate risk level prior to using formal risk assessment, as was found by Perrault and colleagues (2012). A further possibility is that there is some confusion among YPOs regarding the status of youth as high risk or high need, as the YCRNA yields separate ratings for risk (static) and need (dynamic) factors, of which the former are used for the referral criteria. However, comments made on the probation survey would suggest that YPOs are clear on the distinction between the two.

The relationships between the apparent use of the YCRNA to make the referral or to select targets and recidivism were not significant. This was somewhat expected as these practices are removed from the actual program delivery. For example, with respect to the program targets, the specification of needs on the referral may have little bearing on the goals selected for the program by the ISSP worker and even less influence on the actual program activities, which are most likely to contribute to reductions in recidivism.

For the current chapter, some caution is warranted in that the conclusions are primarily based on assumptions made from the data and may not reflect the actual use of risk assessment. YPOs for which there was a YCRNA completed within the month after referral with the same rating as the referral were given the benefit of the doubt and rated as having the best practice present since the forms can remain open for a month, even though YPOs may not have completed the tool until after the referral was submitted. Furthermore, YPOs may have not used the YCRNA in making their referral

but the risk rating could have been the same due to chance, clinical judgment, or other factors. The inclusion of the correct risk rating on the form cannot confirm that the YCRNA was used in the decision-making process to refer to ISSP. The YPO survey arguably provided the better test of whether YPOs use risk assessment to identify youth and formulate program targets, but even YPOs perceptions of whether they use risk assessment may not reflect actual practice, since it appeared that risk assessment was used more often than YPOs reported.

Also, it is possible that there was no relationship between the use of systematic risk assessment for the identification of youth and selection of referral goals due to the failure to take the timing of completion into account for the latter. The use of risk assessment to identify goals variable ratings were based on the selection of targets from the closest available risk assessment, even those completed well after the referral. The correlation analysis was run again including only those cases with YCRNAs on file before the referral, but the relationship was still not significant. The inter-rater reliability for the selection of referral goals was also low, which may have attenuated the relationship between the variables. It was difficult to code the selection of referral goals variable due to the inconsistency of formats and information on the referral forms.

Recommendations. It is important for Youth Justice to introduce guidelines that integrate their risk tool with case management strategies for ISSP. Decision-making about referrals and program targets should be connected to risk assessment, as a properly implemented tool improves decisions (Vincent, Guy, et al., 2012). Guidelines should thus include ensuring that a risk assessment is completed prior to submission of a referral to ISSP. It would be helpful for these guidelines to be developed in consultation with YPOs, as there can be some resistance to adopting a risk tool for decision-making if YPOs perceive a loss of discretion or increases to workload (Ferguson, 2002). For example, in the case where YPOs may wish to refer a low-risk, high-need youth, it may be decided that there are circumstances under which it could be appropriate, as noted in Chapter 3 above with the same cautions.

Another recommendation is to provide clear guidelines to YPOs on the selection of need targets for the referral and goals for the program generally. Criminogenic need targets matched to youths' problematic areas as identified by risk assessment should comprise the majority of targets. Goals should aim to be

comprehensive in addressing youths' multiple needs and the inclusion of non-criminogenic targets or non-problematic need areas should be considered lower priority (Andrews & Bonta, 2010a). Targets should furthermore pertain to specific activities or need areas, as vaguely-worded goals provide limited direction to ISSP workers regarding their roles and duties. While it is understood that the referral constitutes the paperwork for what may be a very detailed case plan discussion with ISSP, it represents the documented information available and should be clear and complete.

Chapter 5.

Best Practices for Contextual Aspects of the Program

Under Blase and Fixsen's (2013) evaluation framework, contextual aspects of the program refer to the populations and settings in which the program is delivered. Since populations are addressed in Chapter 3, this section focuses generally on the systems or contextual aspects of ISSP. ISSP is intended to be delivered primarily in a community setting and, in addition to targeting youths' individual needs, the program aims to address needs or build strengths through the various contexts with which youth interact (Ministry of Children and Family Development, 2006).

Increasingly, the systems that surround youth are being recognized for their important contributions to antisocial behaviour as well as their potential for intervention. Grounded in Bronfenbrenner's ecological systems theory (1979), in which development is framed as occurring within multiple interconnected and interactive systems, antisocial behaviour is viewed as multi-determined with family, school, and community influences in model programs such as Multisystemic Therapy (Henggeler & Sheidow, 2012). Research on the desistance of offending into early adulthood also supports the importance of systems such as family, school, and employment in altering youths' trajectories in this time period (Stouthamer-Loeber, Wei, Loeber, & Masten, 2004).

Unfortunately, justice-involved youth appear to have little success in the very systems that could contribute to their desistance from offending. Engagement may be particularly difficult for offenders whose involvement in these systems is disrupted by time in custody. Custody sentences can separate youth from community resources that they have been accessing and youth may experience delays when attempting to return, such as being unable to re-enrol in school in the middle of the term, losing a treatment bed or other housing placement, having to return to a waitlist for mental health services,

or aging out of the system. British Columbia youth report that short custody sentences are particularly problematic as they disrupt their community life without providing sufficient time to access services and supports in custody (McCreary Centre Society, 2014).

It is thus not surprising that the high school graduation rate for a cohort of British Columbia youth sentenced to custody was 6.2% as compared to the graduation rate of 30% for youth involved in the justice system, which in itself was still much lower than 78% for the general youth population (Representative for Children and Youth & Office of the Provincial Health Officer, 2009). Youth with justice involvement or leaving custody hoping to secure legitimate employment do not appear to fare much better. A study examining offenders' post-release employment found only 30% to be employed in the year after leaving custody, with an equivalent percentage still seeking employment (Bullis & Yovanoff, 2006). Although not all youth come to ISSP from custody, they are still likely to struggle to engage with these systems as the experience of custody does not uniquely account for poor outcomes for justice-involved youth. For example, the difficulties for youth leaving custody of reintegrating into the school system, addressing special education needs, and finding immediately available developmentally-appropriate programs (Anthony, Samples, de Kervor, Ituarte, Lee, & Austin, 2010) could apply to any school disruption due to long-term truancy or drop-out associated with justice-involved youth.

Youth do not come into the youth justice system in a vacuum, nor should individual factors be the sole focus of interventions like ISSP. For the current study, four variables were created to capture the best practices associated with the contextual or systems aspects of ISSP-type programs: an overall systems focus as well as specific best practices relating to the involvement of families, connecting youth to community resources, and a focus on workplace or school re-entry and support. The evidence supporting each of these best practices is reviewed in turn below.

A key feature of many evidence-based youth justice treatment programs is that they have a significant focus on the systems or contexts in which the youth are or may be engaged (Curtis et al., 2004; Mihalic, Fagan, et. al, 2004; Suter & Bruns, 2009). Furthermore, within ISSP-type programs, programs targeting systemic and individual factors using individual treatment yielded the highest average effect sizes in a meta-

analysis of different modalities and targets (James et al., 2013). Findings from meta-analyses of other programs in which youth justice professionals liaise with families and community agencies to provide supervision also reinforce the effectiveness of engaging youths' systems in reducing re-offending (Lipsey, 1999).

Many ISSP-type programs include a specific family component, although the degree of involvement can vary from including families in case planning meetings to offering family therapy. Since many youth will be returning or continuing to live with their families or guardians, it is important that this influential factor not be ignored. Meta-analyses of juvenile justice family intervention programs have demonstrated significant effects for programs that targeted familial supervision/monitoring and affection/communication for recidivism, particularly when the contact was intensive (Dowden & Andrews, 2003; Latimer, Dowden, & Morton-Bourgon, 2003; however see Latimer, 2001). Significant reductions in re-offending, number of arrests, and seriousness of offences for justice-involved youth have also been found for family-based programs (Aos et al., 2001; Curtis et al., 2004). Furthermore, a lack of improvement in family management skills (e.g. supervision, discipline, relationships) was found to mediate the treatment effect for recidivism in an evidence-based systemic treatment program (Eddy & Chamberlain, 2000).

Making connections with community resources may also convey specific benefits for program outcomes. As with family-based approaches, results of meta-analyses for programs providing case management coordination and service brokerage for community programs showed large effect sizes relative to other programs without a community focus (Lipsey, 1999; 2009). In terms of community as a program adjunct, programs with some form of community involvement have demonstrated larger mean effect sizes than those without any community contact (Latimer et al., 2003). Connecting youth to community resources appears particularly important since research suggests that many youth do not seek out services on their own and staff assessments, not adolescents' perceptions of problems, predict service provision for mental health needs (Saunders, Resnick, Hoberman, & Blum, 1994; Stiffman et al., 2000). In their review of re-entry programs, Spencer and Jones-Walker (2004) recommended that intensive transitional supports for housing, employment, and schooling be in place for youth to assist with their successful reintegration. Wiebush and colleagues (2005) also

recommended seeking out community resources to address youths' treatment needs and developing community support networks as part of intensive re-entry programs.

As noted above, school and employment are believed to be potential factors contributing to desistance from a theoretical perspective as well as in empirical research (Sampson & Laub, 2005; Stouthamer-Loeber et al., 2004). In terms of evaluation findings, a juvenile custody re-entry program with similar features to ISSP found educational attainment and employment to be significantly associated with reduced re-offending in young adulthood (Abrams, Terry, & Franke, 2011). Findings from meta-analyses of general youth justice treatment programs indicate that positive changes from these programs in both school participation and vocational accomplishment were associated with decreased re-offending, while changes in psychological factors were not (Latimer et al., 2003; Lipsey, 1995). Wiebush and colleagues (2005) recommended making youths' reintegration into school or the workforce a priority from their evaluation of an intensive re-entry program.

Evaluating program design and delivery. In keeping with best practices, Youth Justice ISSP recognizes the importance of a contextual or systems focus. ISSP workers are required to make weekly contact with families and community agencies (Ministry of Children and Family Development, 2006). There is no ISSP guideline that identifies school or employment re-entry as a priority, although school and employment needs are listed among those that might be targeted in the program.

For this chapter, it is anticipated that the delivery of ISSP will have a systems focus and involve family and community contacts because these are consistent with program guidelines. It is not anticipated that the best practice of prioritizing youth's entry into school or the workforce will be present in a high number of cases. Since all of the variables pertain to the systems supporting youth, it is possible that they may be related to non-recidivism outcomes such as school or employment outcomes, but it is anticipated that these relationships will be strongest for the best practice relating to prioritizing re-entry into school and employment. Likewise, the strongest relationships with the number of programs on ISSP and follow-up are expected for the community and systems best practices. These best practices may be associated with reductions in recidivism, although given that the contextual best practices are primarily aimed at intermediate targets, the benefits may not be seen in the year following ISSP.

Procedures

The contextual variables unfortunately carry some overlap, since it is of interest to determine the contribution of an overall systems approach as well as to investigate the unique contributions of specific systems. An overall systems approach variable was included as it conveys coordination among systems, which may confer additional benefit over individual systems. Furthermore, the systems approach variable required that these contacts were a significant focus of the program. Thus the dichotomous systems variable was defined as having both parent and community contact more months than not. The 5-point variable ranged from having no contact with youth at all to having at least two youth, one parent, and one community contact in a month period, more months than not. The $ICC_{(2,1)}$ value for this variable was 0.87, which exceeded the cut-off for excellent reliability.

The community involvement variable was coded based on any contact that the ISSP worker made or any youth contact that was facilitated by ISSP with schools, substance use and mental health treatment providers, recreational programs, employment programs or jobs, housing services, etc. This definition of community contact differed somewhat from the Youth Justice ISSP definition used in documenting community contact, which refers solely to ISSP worker contact with community agencies. The variables were thus coded from the totals listed by the ISSP worker (where available) and augmented by the notes provided by ISSP and the probation officer. Activities coded under this variable might include rides to school, attendance at integrated case management meetings, and liaising with a substance use counsellor. The dichotomous variable was coded present if any contact had been made with community agencies. The 5-point variable ranged from having no contact during the program with the youth or community agencies to making contact or facilitating youth contact with community resources more often than once a month. The $ICC_{(2,1)}$ value for this variable exceeded the cut-off for excellent reliability at 0.93.

The family contact variable was coded based on the ISSP worker's contact with parents and/or guardians, including group home leaders and foster placement staff. This variable also required referring to the ISSP notes since non-family guardians may not have been included in the Youth Justice ISSP-recorded family contacts. The dichotomous variable was coded based on any ISSP contact with parents or guardians,

while the 5-point variable ranged from no contact with the youth or family to contact made more frequently than monthly over the program. The $ICC_{(2,1)}$ value for this variable was 0.87, which fell in the excellent range.

Regarding the prioritization of youth's entry into school or the workforce, the dichotomous variable was coded present if the number of school- or employment-related activities that ISSP engaged in was comparable to or exceeded the number of other types of activities. The 5-point variable further distinguished qualitatively among types of activities, with those more directly connected to school or work re-entry dictating the highest rating in addition to a focus on these activities. As an example, the ISSP worker taking a youth to register for school, to submit job applications, or providing a ride to an interview were considered to be at a higher level than rides to school or work, helping a youth with a resume, or taking a youth to an employment centre. The $ICC_{(2,1)}$ value for this variable was 0.88, which fell above the cut-off for excellent reliability.

Results

Descriptive data. During ISSP and follow up, 69.1% of youth held a part-time job, while 31.2% were able to obtain full-time employment. Nevertheless, in terms of employment longevity, 46.0% of youth were rated as engaging in no or minimal employment during this time; only 12.1% were mostly employed. With respect to school, 86.9% of youth were enrolled at some point during ISSP and follow up, although 23.4% of youth were rated as minimally or not attending school. Thirty-nine per cent of youth were rated as mostly attending school.

Youth participated in a mean of 2.25 (*SD* = 1.56) programs during ISSP and follow up, with a range of 0 to 9 programs. An average index of missed sessions across community programs for which there were data was also calculated from the limited data available on programs. Approximately 29.5% of youth were rated to have missed a high number of program sessions on average, while 31.8% were rated to have missed a low number of program sessions and the remainder had a moderate level of attendance at programs. Unfortunately, the files did not provide sufficient program data to be included in further analyses other than for the number of programs.

Implementation. The overall systems best practice was rated as present in 24.4% of cases, that is, where contact was regularly made with both family and

community agencies. The individual best practices of any community contact and family involvement were present in 76.7% and 67.6% of cases, respectively. A total of 23.3% of cases had no recorded ISSP or ISSP-facilitated youth contact with community agencies and 32.4% of cases did not record any contact with families or guardians, with 10.8% of cases having no recorded contact with either the community or families. A focus on school or employment was coded in 35.8% of cases, while 40.3% of cases did not show any evidence of activities pertaining to school or employment.

To explore the overlap between variables, the four 5-point variables were correlated (see Table 13). All of the variables showed between a moderate to large correlation effect size with each other. As such, they were analysed separately in the following sections.

Outcomes. The relationship between contextual variables and outcomes including recidivism were examined using hierarchical logistic regression for dichotomous outcomes and hierarchical linear regression for continuous file outcomes. Covariates for the non-recidivism outcomes were identified using correlation analyses (See Table 14). Variables indicating whether youth were employed or enrolled in school at the beginning of ISSP were included for relevant analyses that did not reflect a dichotomous absent/present outcome. Due to the significant correlations with best practice predictors, interaction terms were created for age and a systems focus (r = -0.35, $p < 0.001$), community involvement (r = -0.32, $p < 0.001$), and family contact (r = -0.41, $p < 0.001$). The variables were centered prior to creating the interaction terms due to lack of a meaningful zero point for age and the problems with multicollinearity introduced by the terms (e.g., inflated standard error). A file time variable was created for time at risk to reflect the number of months for which there was file information during the follow-up period as well as the time spent in the community during ISSP and follow up. The file time variable was significantly correlated with each of the target individual and systemic factors (r = 0.34, $p < 0.001$), community involvement (r = 0.36, $p < 0.001$), family contact (r = 0.30, $p < 0.001$), and prioritize school/employment (r = 0.41, $p < 0.001$) variables, such that interaction terms were created for the analyses. Interaction terms were also created for file outcome quality and community involvement (r = 0.18, $p < 0.05$), file outcome quality and prioritizing school/employment re-entry (r = 0.19, $p < 0.05$), youth referred from the Victoria Custody

Centre and family involvement (r = -0.16, p < 0.05), and addictions and prioritizing school/employment re-entry (r = 0.16, p < 0.05) that were included in moderator analyses where both the covariate and contextual variables were present.

For employment outcomes, obtaining any part-time employment during ISSP and the follow-up period was significantly associated with community involvement (B = 0.36, SE(B) = 0.18, Wald $\chi^2(1)$ = 3.93, Exp(B) = 1.43, p < 0.05) prior to the addition of the interaction term to the model, which was not significant. A systems focus, family contact, and prioritizing school/employment were not significantly associated with part-time employment (see Tables 15-18). Any full-time employment during ISSP and follow up and youths' overall employment during the study period were not significantly related to any of the contextual variables (see Tables 19-26). Youths' employment status at the end of ISSP was significantly related to a priority on school or employment activities, B = 0.47, SE(B) = 0.21, Wald $\chi^2(1)$ = 4.82, Exp(B) = 1.60, p < 0.05, prior to the addition of interaction terms to the model, although the interaction terms were not significant. Employment status at the end of ISSP was not associated with any of the other contextual variables (see Tables 27-30).

Regarding school outcomes, none of the contextual variables were significantly associated with any school attendance (see Tables 31-34). Youths' overall level of school enrollment during ISSP and follow up was significantly associated with a systemic focus (β = 0.13, SE(B) = 0.05, t(154) = 2.60, p < 0.05), community involvement (β = 0.12, SE(B) = 0.04, t(154)= 2.00, p < 0.05), family contact (β = 0.15, SE(B) = 0.04, t(154)= 2.33, p < 0.05), and school/employment priority (β = 0.12, SE(B) = 0.04, t(154) = 2.05, p < 0.05). None of the interaction terms were significant when added to the models (see Tables 35-38 for the full models). The contextual variables were not significantly related to school enrollment at the end of the follow-up period (see Tables 39-42).

The number of programs that youth participated in during ISSP and the follow-up year was considered an outcome variable, as ISSP could facilitate referrals to or participation in such programs. The contextual variables were not significantly associated with the number of programs (see Tables 43-46).

In terms of any recidivism, none of the contextual variables significantly predicted recidivism, although the relationships were all in the anticipated negative direction (see Tables 47-50).

Discussion

A majority of Youth Justice ISSP workers made contact with community agencies or facilitated youths' contact with community agencies during the program, as well as made contact with families. However, only one-quarter of files were rated as having significant focus on the systems in which youth were engaged. A larger proportion of ISSP workers than anticipated (approximately one-third) prioritized school or employment activities, given that a school/employment focus is not an ISSP guideline. Nonetheless, significant proportions of files had no documented contact with community agencies, families, or school. Furthermore, it should be noted that even the highest category of the data-driven 5-point variables fell short of the current Youth Justice ISSP guidelines regarding contacts for families and community agencies, requiring that the contact guidelines for a week period instead be met in a month. Given that ISSP is intended to provide support for a period of approximately six months, it is critical that this time is used to engage participants' community and family supports so that the benefits of ISSP may continue well beyond the end of the program. Regardless of whether youth are referred to ISSP in custody or in the community, a significant focus of ISSP should be facilitating connections with community agencies and ensuring that when ISSP finishes there will be community resources to take over. Since ISSP targets high-risk youth, it is unrealistic to imagine that at the end of ISSP participants' criminogenic needs will have resolved or that ISSP alone will be sufficient to address these needs.

Contextual variables were significantly associated with a number of intermediate outcomes, such as part-time employment during ISSP and follow up, employment at the end of ISSP, and overall school enrollment. Generally speaking, they were in the anticipated direction and stronger for the most theoretically-relevant variables. An exception was the stronger than anticipated relationship between family involvement and school enrollment, which was comparable to the effect sizes for community involvement and prioritizing school/employment re-entry. Although a number of other relationships with school or employment outcomes were not significant or only marginally significant,

these also tended to be positively related to outcomes and strongest for conceptually-related contextual variables.

There were no significant relationships between contextual variables and recidivism, although the effect sizes were generally in the anticipated negative direction (i.e., reducing the likelihood of recidivism) other than the emphasis on school and employment re-engagement variable. Some research suggests that part-time employment may increase the risk of offending (Brame, Bushway, & Paternoster, 2004), although others argue that there are several mediating factors such that the positive relationship between employment and offending may be spurious (e.g., Staff, Osgood, Schulenberg, Bachman, & Messersmith, 2010).

In terms of limitations of this chapter, the inclusion of time at risk during ISSP and follow up for the outcome analyses is likely to have attenuated the effect of the best practices due to the overlap in variance (see Miller & Chapman, 2001). Also, although attempts were made to control for file outcome quality, for which better data quality could lead to higher ratings for contextual variables as well as to the detection of positive outcomes, it is still possible that data quality contributed to some of the shared variance between contextual and outcome variables. However, the fact that the relationships between contextual variables and outcomes were generally stronger for variables for which an association was expected suggests that some variance in the relationship was independent of time at risk and file quality. Another difficulty with the contextual variables pertained to the overlap between these items. Due to their shared variance, it is difficult to parse out the contributions of individual variables. Although there was conceptual overlap between some variables (e.g., community and family contact defining a systems approach), the variables were nonetheless developed to reflect distinct best practices.

Most importantly, the retrospective nature of the study cannot clarify the direction of the relationships between the variables. For example, while overall enrollment in school during ISSP and follow up was treated as an outcome, it is equally likely that youth enrolled in school were more likely to have community involvement on ISSP due to the availability of a school placement to make community contacts. The direction of significant relationships for the employment variables is similarly unclear. Furthermore, it has been noted in desistance research that, to some extent, the relationship between

recidivism and school or employment outcomes may reflect an overall pattern of improved adjustment rather than a causal relationship (Stouthamer-Loeber et al., 2004).

For the family involvement best practice in particular, it appears likely that a lack of clear documentation contributed to the relatively poor implementation recorded in the current study. However, it may also be that it is difficult for ISSP workers to engage parents despite their best efforts; the McCreary Centre Society report (2005) on youth in custody indicated that nearly half of Burnaby Custody Centre youth reported that their parents never or almost never visited them. Reasons for these difficulties may include practical barriers (e.g., parents being at work during ISSP hours), alienation from youth, or that parents may be reluctant to engage in the program because they feel blamed for their children's behaviour (Kumpfer, 1993; Walker & Friedman, 2001).

Recommendations. Caution is necessary in making recommendations for these best practices without further information on the reasons for their apparently poor implementation according to the ISSP guidelines. If contacts were occurring but were not recorded, as often appeared to be the case, it is important that documentation of these contacts is improved. Another possibility is that the guidelines are unrealistic; for example, if ISSP workers have 10 youth on their caseload and are expected to spend a minimum of two hours per youth per week in the program, presumably in addition to paperwork and travel time, weekly community and family contacts may not be viable. These guidelines may need to be revised with input from ISSP staff.

If the level of contact indicated by the guidelines is feasible, there are several recommendations that may assist with the implementation of the best practices. It may be helpful, for example, if training were to include a brief summary of systems theory, as it is important for staff to understand the theoretical basis for a program (Gendreau, Goggin, & Smith, 1999), or in this case, the best practices. Scheduling regular intensive case management (ICM) meetings and ensuring that ISSP is included may be another way to build greater family and community connections. It was often noted in the file that ISSP was not able to be present at ICMs due to meetings occurring on their days off, so it may be useful to have a process in place when scheduling ICMs to overlap with ISSP days through better coordination, setting schedules farther in advance, etc. Although some ISSP staff had an alternate ISSP worker attend in their place, it would generally be best for building relationships if it is the participant's primary ISSP worker who attends.

Changes might also be made to the program to encourage more participation from families, including consulting families in setting the goals for ISSP and keeping them informed of the ISSP process and of the service options (Walker & Friedman, 2001). Some ISSP workers documented their introduction session with families in the file, suggesting that some families are already being informed of the ISSP process. It would be helpful for these introductory sessions to be consistently documented to study the relationship with family engagement later on in the program. It may also be useful to solicit feedback from families to ensure that they feel approached in a respectful and nonjudgmental manner and that their authority is supported in the presence of ISSP, which may contribute to greater parental participation (Walker & Friedman, 2001). Furthermore, if it is impractical to engage parents or they are unwilling, it may be helpful to consider engaging extended family (Brock, Burrell, & Tulipano, 2006).

To increase the level of community focus, it may be helpful that the role of ISSP as facilitating community involvement is emphasized in training, such that activities that engage youth in community resources are prioritized over individual activities (e.g., going for coffee). Since treatment providers' knowledge of community resources is predictive of the number of services accessed by youth (Stiffman et al., 2000), it is important that ISSP workers are well-informed about the services available. As community agencies are often in flux and keeping up with the changes in available resources would take valuable time away from youth, a community resource coordinator position ideally would be created to develop a multiagency service network and service guide for ISSP workers that would be regularly updated (Lattimore et al., 2004).

In terms of the prioritizing school and work best practice, it may be helpful to include this practice in the program guidelines for Youth Justice ISSP since it was associated with outcomes in the current study. However, it may be modified to pertain specifically to youth for which school and employment are criminogenic needs, since a focus on these needs may take time and resources away from other areas in greater need of intervention if youth are already engaged in these activities. Also, it may be important to maintain the idea that the best practice pertains to re-entry rather than being interpreted as it was in the current study as any time spent engaged in school or employment activities. For example, time spent driving a youth to school who is already regularly attending may be better directed at helping the unemployed youth find a job.

Chapter 6.

Best Practices for Structural Elements of the Program

Structural elements of a program define the boundaries within which the program activities operate. Although structural elements are often a minor consideration in evaluation, research on juvenile justice programming indicates that it is not just the type of intervention that participants receive but how much that dictates successful outcomes (Lipsey et al., 2010). The amount of intervention, or dosage, might refer to the length, frequency, or number of sessions, as well as the total duration or length of the program (Nation et al., 2003). Structural elements of programs can also encompass aspects such as a group format or a low number of youth on a caseload (Blase & Fixsen, 2013); however, since these features remained relatively consistent over participants or the necessary data could not be retrieved from files, they were not included in the study.

Meta-analyses of youth justice programs indicate that effective programs tend to be more intensive, either in terms of hours of weekly contact and/or total hours of service (Lipsey, 1992; 1995; 1999). The recommendations emerging from meta-analysis are that programs should exceed 100 hours in twice-weekly contacts averaging five hours per week (Lipsey, 1995). However, other meta-analyses have found no effect of program intensity on recidivism or even for intensity to be negatively associated with effect size in the case of low-risk youth (Andrews & Dowden, 2005; Latimer et al., 2003; Lipsey, 2009). A meta-analysis of intensive re-entry programs found that higher intensity, as measured by number of contacts per month, yielded larger effect sizes (James et al., 2013). It is thus possible that research on general juvenile justice programming is only able to provide limited direction for ISSP best practices since intensity may be moderated by program type.

Meta-analytic findings for the importance of duration for juvenile justice programs are similarly equivocal. Meta-analyses have generally found that programs of longer duration, using median lengths of 18 and 28 weeks, are positively associated with reductions in recidivism (Lipsey, 1992; 1995; 1999; but see also Lipsey, 2009). However, Latimer and colleagues (2003) found that programs exceeding six months yielded smaller effect sizes than shorter programs. Unlike the findings for intensity in the study, the results for duration were consistent across risk levels.

Evaluations of ISSP-type programs are not able to provide any more clarity on the issue of duration. A meta-analysis of intensive re-entry programs did not find duration to be significantly related to effect sizes for recidivism (James et al., 2013). However, an evaluation of a custody re-entry program with similar features to ISSP found that the average length of the program for participants who did not re-offend during the follow-up period was significantly longer by 2.3 months (Abrams et al., 2011). It should be noted that there was significant selection bias in the sample and it was not clear if longer program duration was associated with a third variable (e.g., youth characteristics, risk) that mediated the relationship between program duration and offending. Generally speaking, researchers in the area of ISSP-type programs assert that the duration of the program is important, regardless of the number of sessions involved (Gies, 2003). In their review of effective program characteristics for youth re-entering the community, Spencer and Jones-Walker (2004) recommended that youth participate in re-entry programs for at least nine months.

There are several factors contributing to the lack of clear findings for dosage for youth justice programs. Structural elements often appear to be overlooked, as both individual studies and meta-analyses attempting to include these variables have noted a failure of the information to be reliably recorded and reported in studies, respectively (Lipsey et al., 2010; Suter & Bruns, 2009; Unruh, Gau, & Waintrup, 2009; Vieira, Skilling, & Peterson-Badali, 2009). For example, Lipsey (2009) indicated that the coding of intensity for the meta-analysis often had to be based on the *intended* intensity of programs rather than their actual intensity due to an absence of information for more precise coding. Furthermore, the Abrams and colleagues (2011) evaluation study suggests that the effects of duration may be masked in meta-analyses due to significant within-program variation. In the case of their evaluation, though the average duration

that might have been recorded for the program in a meta-analysis would be 8.3 months, there was a range of one to 15 months for youth. As such, the information available from meta-analysis currently may not be the best way to capture the effects of these variables until more implementation data are available and analysed within studies. Clearly, the equivocal findings and quality of evidence for the best practices of intensity and duration indicate a need to empirically examine dosage in the context of ISSP.

A further issue related to treatment intensity and duration is how the risk principle of the RNR model (Andrews et al., 1990) might be applied within ISSP. Consistent with the risk principle, higher-risk youth should receive a higher-intensity program than low-risk youth as high-risk offenders stand to benefit the most from treatment (Andrews & Bonta, 2010a). In fact, Latimer and colleagues (2003) recommended limiting programs to six months and a maximum dosage of 20 hours for low-risk offenders. Listwan, Cullen, and Latessa (2006) emphasized the importance of not having a fixed duration or intensity for re-entry programs, but tailoring these to youth based on an assessment of their risks and needs.

However, in practice, there may be pressure or a greater inclination for juvenile justice staff to focus on low-risk youth because they are more likely to be cooperative and motivated (Bonta & Andrews, 2007). The term YAVIS (young, anxious, verbal, intelligent, and social) has been coined for these preferred clients, while high-risk youth may be avoided due to perceptions that they are hardened and less amenable to treatment (Lipsey & Wilson, 1998). High-risk youth may also present practical barriers for a more intensive level of ISSP due to problems with unstable housing, difficulties maintaining contact information, avoidance of ISSP due to outstanding warrants, and disruption or premature termination of the program during periods of custody.

Evaluating program design and delivery. Regarding the correspondence between Youth Justice ISSP and best practices, the guidelines stipulate that ISSP workers are expected to dedicate a minimum of two hours per week to each youth (Ministry of Children and Family Development, 2006), with no guideline pertaining to total contact hours. The guidelines indicate that youths' participation is reviewed after every six months of ISSP and there is an option to extend program length based on a re-assessment of youths' risk level and availability considerations. In terms of the risk principle within the program, there is no distinction for dosage based on risk made in the

Youth Justice ISSP guidelines, though given that the program is targeted at high-risk youth, further dosage specifications based on risk level may not have been seen as necessary. The ISSP guidelines recommend flexibility around intensity and suggest that more intensive service may be needed at the beginning of the program (e.g., five hours per week), which appears to reflect the assumption that risk should decrease as youth continue in the program.

For the current chapter, the implementation hypotheses of interest are whether youth in the program received a sufficient dosage according to the best practices for the total number of program hours and duration, as well as according to the Youth Justice ISSP guidelines for monthly contact hours and duration. It is anticipated that, consistent with much of the previous research on dosage (James et al., 2013; Lipsey, 1992; 1995; 1999), intensity and duration will be significantly associated with recidivism in the year following the program. Furthermore, it is of interest to determine if risk was positively associated with dosage, although it is anticipated that there will be no relationship or a negative relationship since there are no program guidelines pertaining to adjusting dosage based on risk and due to the barriers to engaging high-risk youth.

Procedures

Two separate variables for intensity were created due to the differences between the best practice and Youth Justice ISSP guideline. As with other variables, the Youth Justice guidelines were included as it may be helpful to maintain current guidelines if adherence to the best practices does not confer any additional benefit. The best practice for intensity was coded based on the total program hours, because ISSP hours were recorded by month rather than weekly and the five hour weekly contact best practice may not be as suitable for an ISSP-type program of longer duration as to other types of programs. The dichotomous best practice intensity variable was coded based on whether or not the total hours exceeded Lipsey's (1995) recommended average of 100 hours. The 5-point variable was coded based on the distribution of hours for the program, ranging from zero to greater than 150 hours of program service. The actual hours calculated from the files were also used in the analyses since the information was available and coding the number of hours into a 5-point variable could serve to attenuate

the effects of intensity. The $ICC_{(2,1)}$ value for this variable was 0.97, falling above the cut-off for excellent reliability.

For Youth Justice ISSP intensity, the dichotomous variable was based on the guidelines stipulating that two hours should be spent on a youth per week, such that it was coded present if ISSP workers recorded eight hours of service time more months than not. The 5-point variable was coded based on varying levels of adherence to the guidelines occurring in practice, with further distinctions made for the amount of direct contact with youth (as opposed to paperwork, probation meetings, etc.). The variable ranged from no contact with youth noted to at least eight hours of direct contact (i.e., in-person with youth) for at least one third of the program. The $ICC_{(2,1)}$ value for the Youth Justice ISSP variable also exceeded the cut-off for excellent reliability at 0.82.

Two separate dichotomous variables were also created for duration due to the differences between the best practice and the Youth Justice ISSP guideline. For the best practice duration variable, it was coded present if youth participated in the program for equal to or greater than nine months. For the Youth Justice ISSP duration variable, youth were coded as having the guideline present if they participated in the program for equal to or greater than six months. A single 5-point variable for duration for the best practice and ISSP guidelines was created based on the frequency distribution, which ranged from zero months in the program (i.e., no contact noted in the file) to greater than 24 months in the program. The variable was calculated on the actual program duration obtained from the file (instead of the database) and included months spent in custody, in residential treatment programs, etc. The $ICC_{(2,1)}$ value for the duration variable was 0.75, which fell at the cut-off for excellent reliability. As with intensity, since continuous data were available for the number of months that youth participated in the program, actual duration was included in a separate analysis in the event that the coding of duration into a 5-point variable attenuated its effect. Given that programs could be disrupted or the activities restricted during youths' time in custody, long periods of absence, or while in non-local residential programs, thereby changing the nature of ISSP as a community program, a separate index of community time was calculated and included in the analysis.

Results

Descriptive data. Youth received a mean of 44.80 (*SD* = 61.35) program hours, with a range of 0 to 296 hours. Youth participated in a mean of 11.77 (*SD* = 7.48) months of ISSP, with a range of 0 to 34 months. The mean number of months youth participated in ISSP in the community was 10.21 (*SD* = 7.22), while the range was the same as the total program duration.

Implementation. In accordance with best practice research suggesting a minimum of 100 program hours, 14.8% of youth exceeded this cut-off. For the Youth Justice ISSP guidelines, 16.5% of youth were rated as having the intensity guideline present; that is, they received a minimum of eight program hours per month, more months than not. At the same time, 33.5% of youth were not recorded as receiving the eight hour minimum of programming at any time while in ISSP.

In terms of program duration, 59.1% of youth received at least nine months of programming as recommended by the best practice. The Youth Justice ISSP guideline of six months or more of programming was met in 69.9% of cases.

Outcomes. Logistic regression analyses were used to examine the relationship between structural variables and re-offending in the year following ISSP after controlling for the covariates identified in the methods section. Low cognitive functioning was correlated with the best practice intensity variable ($r = 0.17$, $p < 0.05$), Youth Justice ISSP intensity variable ($r = 0.17$, $p < 0.05$), raw total hours ($r = 0.17$, $p < 0.05$), and months on ISSP ($r = 0.16$, $p < 0.05$) so interaction terms were created for moderation analyses.

The best practice intensity variable marginally predicted recidivism prior to the addition of the interaction term to the model ($B = -0.30$, SE(B) = 0.18, Wald $\chi^2(1) = 2.72$, Exp(B) = 0.74, $p = 0.10$). The interaction term with low cognitive functioning was also marginally significant when added to the model. Further analyses were conducted using general linear regression to derive and plot the marginal means for recidivism for low cognitive functioning and intensity, revealing that intensity led to greater decreases in recidivism for youth with low cognitive functioning than for youth without identified cognitive functioning issues (see Figure 3). The Youth Justice intensity and the duration

variables were not significantly related to recidivism. See Tables 51 to 53 for the complete models.

Since raw continuous scores were available for the intensity and duration variables, these were tested as alternate dosage indices. Raw total hours of ISSP significantly predicted recidivism (B = -0.01, SE(B) = 0.00, Wald $\chi^2(1)$ = 5.50, Exp(B) = 0.99, $p < 0.05$) prior to the addition of the interaction term to the model. Youth documented as participating in 10 hours of ISSP thus had an odds ratio of recidivism of 0.92, while youth participating in 100 hours of ISSP had an odds ratio of 0.44. The interaction term for low cognitive functioning and number of hours was also significant (B = -0.01, SE(B) = 0.01, Wald $\chi^2(1)$ = 3.02, Exp(B) = 0.99, $p < 0.01$). See Table 54 for the complete model. Due to the continuous nature of the variable and the number of levels, marginal means were not plotted. The analysis for hours and recidivism was run separately for youth with and without low cognitive functioning with only the SAVRY total score in the model, as the other variables caused problems with multicollinearity. The relationship between number of hours and recidivism was significant for youth with low cognitive functioning only (see Table 55). Youth with low cognitive functioning participating in 10 hours of programming had an odds ratio of recidivism of 0.80, while those participating in 100 hours of programming had an odds ratio of 0.11.

The total number of months that youth participated in ISSP did not predict recidivism (see Table 56). However, the number of months that youth participated in ISSP while in the community (i.e., not while in custody or in a non-local treatment program) was significantly related to recidivism (B = -0.06, SE(B) = 0.02, Wald $\chi^2(1)$ = 4.91, Exp(B) = 0.95, $p < 0.05$). See Table 57 for the complete model. Youth participating in ISSP for a year period had an odds ratio of recidivism of 0.52.

Independent samples t-tests were used to determine whether youth rated as high risk and meeting program selection guidelines using the dichotomous variables in Chapter 3 received more programming, respectively. Youth rated high risk did not receive more hours of programming, $t(42.60) = 1.98$, $p = 0.05$, equal variances not assumed, or a program of longer duration, $t(168) = 1.20$, *ns*, equal variances assumed, than low-risk youth, although the former approached significance. Participants who met Youth Justice ISSP criteria did not receive more hours of programming, $t(172) = 1.58$,

ns, equal variances assumed, but remained on ISSP significantly longer, $t(172) = 2.50$, $p < 0.01$, equal variances assumed, than youth not meeting criteria.

Discussion

In terms of the implementation of the program, the average recorded intensity fell well below the 100-hour best practice at 45 hours, with only around 15% of youth exceeding the best practice for intensity. Approximately 16% of youth received eight or more program hours per month, more months than not, which is consistent with Youth Justice ISSP guidelines. Though ISSP workers may be unaware of the best practice for intensity, if youth had received eight hours of contact per month according to the ISSP intensity guidelines and participated in ISSP for an average duration of one year, youth should have received approximately 80-100 hours of programming, consistent with best practice. The discrepancy between the guidelines and practice may reflect conservative estimates of hours from the incomplete data provided, although a number of files for which most logs were completed still reflected fewer hours than those stipulated by the guidelines.

With respect to duration, youth participated in ISSP for an average of nearly one year, while the average number of months actually spent in the community was approximately 10. These both exceed the best practice of nine months for duration, which in practice occurred for nearly 60% of youth. It also appears that the six month guideline for ISSP was routinely extended in practice, since the program duration exceeded six months for almost 70% of youth. Long ISSP terms appeared to be particularly common for youth who committed sexual offences, who often participated in ISSP for significant portions of their probation order to supplement the supervision provided by the YPO.

The coded intensity and duration variables did not show a significant relationship with offending, but the relationships were in the anticipated direction of being negatively associated with recidivism. The raw hours and number of months in the program showed an interaction with cognitive functioning, such that youth with low cognitive functioning appearing to have benefitted more from greater intensity and duration in terms of reduced recidivism. The actual number of months that youth spent in the community participating in ISSP was also negatively associated with recidivism. Of note

was the discrepancy between the significant finding for recidivism for the actual months that youth spent engaged in ISSP in the community and the non-significant finding for the entire program length. This discrepancy could help to explain the equivocal findings for duration in meta-analyses (James et al., 2013; Lipsey, 2009), as the average program duration or estimated program length indices that are often used in meta-analysis may not provide an adequately precise or accurate measure of duration.

Youth rated as high risk did not receive significantly more months of programming but approached significance for a higher number of hours than low-risk youth. Given the minimal direction from ISSP guidelines around modulating intensity and duration to risk level and the likely difficulties of engaging high-risk youth, the findings of a minimal effect for risk and dosage were not surprising. The general trend towards providing more intensive service to the most appropriate youth was nevertheless encouraging. Whether or not participants met Youth Justice ISSP guidelines was significantly associated with the number of months in the program. This reflects, to some extent, the recognition of the risk principle in that the ISSP guidelines are meant to identify high-risk youth.

With respect to limitations for this chapter, a reliable index for hours was often lacking from the files, either due to missing logs or information missing from logs (e.g., logs formatted to report number of contacts rather than number of hours). Many of the total program hours had to be estimated from probation notes or the number of occasions documented, such that considerable error was likely to have been introduced and to have attenuated the effects of the variable. Probation file entries such as "meeting weekly with ISSP" provided little direction regarding how many hours youth had participated in ISSP. Furthermore, it was impossible to separate program duration from other factors, such as maturity, that may have contributed to the significant findings for these variables. The presence of a significant relationship for community time versus total time could suggest a unique effect of the program, although this variable may also be confounded with youth characteristics, negative effects yielded from time spent in custody, or other factors.

Recommendations. It may be beneficial for Youth Justice to review the guidelines pertaining to program contact and duration. As with the community and family contact guidelines, two hours weekly per youth may reflect an unreasonable

expectation with a caseload of ten youth. While the guidelines indicate that, with a caseload of a maximum of 10 youth, ISSP workers should be able to devote approximately 3.5 hours to each youth (Ministry of Children and Family Development, 2006), these ideals may not take into account travel time, paperwork, liaising with YPOs, time spent tracking down youth or no shows, and difficulties organizing time around the schedules of 10 youth. It would be helpful to solicit feedback from previous and current ISSP workers regarding ideal caseload sizes and intensity guidelines or to consider staffing changes since best practices would ideally drive the ISSP guidelines rather than practical limitations. It may be possible through some combination of changes to weekly intensity and duration guidelines to achieve the best practice for total program hours without making significant changes to the workload of ISSP staff.

An important recommendation relates to the risk principle for intensity of treatment. Although ISSP is intended for high-risk youth, it is clear from Chapter 3 that low-risk youth are referred to and participate in the program. Drawing on the findings from Latimer and colleagues (2003), it may be beneficial to add more specific guidelines to ISSP relating to risk prioritization within the program. It may be beneficial for training for ISSP workers to include the RNR model (Andrews et al., 1990) and the negative outcomes that can be associated with involving low-risk youth in justice programming. Consistent with Latimer and colleagues' (2003) recommendation, it may be helpful to place a limit on the duration and intensity for low-risk youth (e.g., six months, 20 hours).

Generally speaking, the results of the current study suggest that youth benefit from programs of a longer duration and that youth with low cognitive functioning particularly benefitted from greater intensity. However, due to the significant resources necessary to provide a higher number of hours or to retain youth in ISSP for a longer duration, it would be beneficial to further examine these best practices in a prospective study before making changes to the current program guidelines.

Chapter 7.

Best Practices for Program Activities

Though structural elements are clearly an important facet of ISSP, there are still others who maintain that, assuming a minimum dosage, it is as much the quality of contacts as the quantity that make the difference in ISSP-type programs (Altschuler & Armstrong, 2004). Particularly for programs like Youth Justice ISSP that are relatively unstructured compared to other evidence-based youth justice programs, the types of activities and composition of activity types are likely to vary widely. Best practices are important in providing direction to ISSP workers regarding their selection of program activities. Unfortunately, much of the evaluation research on similar programs to ISSP does not provide specific analysis of the program activities or explore the relationship between types of activities and recidivism. The lack of implementation data on program activities is also likely to contribute to the significant variability found in effect sizes and the null findings in the evaluation studies of ISSP-type programs.

The need principle of the RNR model (Andrews et al., 1990) indicates that, where the goal of a program is to reduce recidivism, it is important that the program's activities target youths' criminogenic needs, or needs that are functionally related to their offending behaviour. Andrews and Bonta (2010a) identified a "Big Four" group of risk factors that appear to show the strongest relationship with re-offence risk in juveniles and adults in meta-analysis research. Another four risk factors demonstrating a moderate relationship with re-offending complete the "Central Eight." The former are a history of antisocial behaviour, an antisocial personality pattern (e.g., anger problems, impulsivity), antisocial cognition or attitudes, and antisocial peers, while the latter are family circumstances, school/work engagement, leisure/recreation, and substance abuse. Andrews and Bonta (2010b) further delineated non-criminogenic needs, which are not generally associated with reductions in offending, including self-esteem, vague emotional problems, and lack of physical activity. As such, and is evident from the

multiple duties listed in the ISSP guidelines (Ministry of Children and Family Development, 2006), ISSP workers have a wide range of options when choosing activities for the program.

Meta-analyses of youth and adult justice programs have consistently supported that programs are most effective at reducing re-offending when activities target at least one of the criminogenic needs listed above (Andrews & Bonta, 2010a; Dowden & Andrews, 1999; Gendreau, 1996). However, researchers have extended the need principle to indicate that programs should target youths' *multiple* criminogenic needs (Andrews, Bonta, & Wormith, 2006; Mackenzie, 2006). Latessa and Lowenkamp (2005) recommend that programs target at least four to six criminogenic needs to yield significant reductions in recidivism. ISSP is ideally suited to address this best practice due to the broad focus on a number of need targets.

However, one limitation of much of the intervention literature supporting the RNR model is that meta-analysis has focused generally on criminogenic needs without being able to speak to the effects that these principles may have on individual service delivery when directed at youths' specific criminogenic needs (Hoge, 2002). In a recent meta-analysis of 374 tests of intervention effects from the Carleton University databank, Andrews and Bonta (2010a) found that programs that involved individualized need matching were associated with significantly larger effect sizes. Unfortunately, no further information was available to determine whether needs matching was a feature within the programs or whether this finding referred to the individuals being matched to one or more specific programs for which they had been identified as having a need.

Vieira and colleagues (2009) conducted a study in which they found a significant reduction in re-offending when youth were matched to programs that were appropriate to their specific criminogenic needs as identified using a risk assessment tool. In fact, the relationship between the percentage of successfully matched criminogenic needs and re-offending was stronger than for the overall number of programs youth received that targeted criminogenic needs. Interestingly, the relationship between needs match and re-offending was not significant for females when equal-sized groups of female and male youth were analyzed separately, despite the two groups being otherwise equivalent on the number of needs identified by clinicians and the percentage of needs successfully matched to programs (Vitopoulos, Peterson-Badali, & Skilling, 2012).

While the previous studies pertained to needs matching across different treatment programs as directed by probation, other evidence-based youth justice programs like Multisystemic Therapy and Wraparound services include an individualized focus as a core component within the program itself (Bruns & Walker, 2008; Henggeler, 1999). Moreover, five of six ISSP-type programs had individualized or specific case planning of activities as a core program characteristic in a review of promising re-entry programs (Gies, 2003).

A final consideration with respect to program activities is the balance of the types of activities in the overall program, specifically with respect to support and supervision activities. In a revised discussion of the RNR principles, Andrews and colleagues (2006) commented that there may be several compelling reasons to include non-criminogenic need targets in youth justice programs. In particular, given the nature of ISSP as serving a mentoring function, it may not be reasonable to expect that all activities would pertain directly to criminogenic needs. For example, it may help to establish and maintain the relationship with youth if staff include reward activities recognizing youths' achievements or spend some time engaged in informal, non-criminogenic need activities. Even so, several researchers have emphasized the importance of ensuring that criminogenic needs remain the focus of programs (Andrews et al., 2006; Latessa & Lowenkamp, 2005; Mackenzie, 2006). A meta-analysis of youth and adult justice programs found that the effect size for recidivism was strongly correlated with programs that targeted more criminogenic needs than non-criminogenic needs (Andrews & Bonta, 2010a).

In addition to non-criminogenic need activities, ISSP-type programs serve an important function in providing supervision for youth who might otherwise be placed in custody. However, the supervision aspect of ISSP should not dominate the program either. In the adult literature, intensive supervision programs that featured a treatment component reduced recidivism by 10% relative to programs that solely comprised surveillance measures (Gendreau, Goggin, Cullen, & Andrews, 2000). For youth, a critical component of most ISSP-type programs is that along with supervision, there is a significant focus on treatment (Gies, 2003). The Intensive Aftercare Program model requires a balanced mix of targeting risk (i.e., through supervision measures) and providing interventions aimed at offenders' criminogenic needs (Altschuler & Armstrong, 1994). Thus, a significant treatment focus appears to be important in reducing

recidivism, rather than a predominant emphasis on supervision or non-criminogenic need activities.

Evaluating program design and delivery. For Youth Justice ISSP, activities targeting criminogenic needs comprise a number of the duties of the ISSP worker under the program guidelines (Ministry of Children and Family Development, 2006). The guidelines reference criminogenic needs generally and indicate that ISSP should support activities and programs targeted at youths' needs. However, explicit guidelines regarding what constitutes a criminogenic need, a minimum number of criminogenic need targets, or the optimal balance of criminogenic, supervision, and non-criminogenic need activities are lacking.

For the implementation of these best practices, some variation was anticipated in the number of criminogenic need activities, the number of specific need activities, and in the balance between support and supervision activities due to a lack of ISSP guidelines for these. Individual ISSP workers are likely to vary in their familiarity with research on the RNR model without specific training or guidelines through ISSP. As in Vieira and colleagues (2009), it was of interest to determine whether particular needs were commonly unaddressed as it may indicate a training gap.

Consistent with previous research, it was expected that youth who participated in activities targeting multiple criminogenic needs would be less likely to re-offend than those participating in non-criminogenic need activities. Including individual best practices for targeting youths' specific needs and engaging in more criminogenic need activities than non-criminogenic needs or supervision activities allowed for the examination of whether these practices also conferred any benefits for reductions in recidivism.

Procedures

Examples of activities that were coded as targeting criminogenic needs are included in the coding manual that is available upon request from the author. All needs included on the YCRNA were considered criminogenic for the purposes of the current study, even though housing is not considered one of the primary criminogenic needs in the general literature (Andrews & Bonta, 2010a). Nevertheless, though it is not clear what risk mechanisms underlie the relationship, the number of out-of-home placements

is a risk factor for youth offending (Cottle, Lee, & Heilbrun, 2001). It may be that housing needs are more prominent for youth, who have fewer resources and skills to seek their own housing than adults. Furthermore, one might argue for the importance of including housing as a target based on a hierarchy of needs, where the significance of school and treatment appointment attendance is generally outweighed by the urgency of addressing survival needs.

The coding for the criminogenic needs variables was based on Latessa and Lowenkamp's (2005) recommendation that programs should target between four and six needs. Thus, the dichotomous variable was coded based on the presence of activities supporting at least four criminogenic needs. The 5-point variable ranged from no contact with the youth to multiple activities targeting at least four needs. The $ICC_{(2,1)}$ value for the variable was 0.88, which exceeded the cut-off for excellent reliability.

The coding for the specific needs match variable relied on the referral YCRNA or closest YCRNA to start date to identify needs rated as being at least moderately in need of intervention. Where ISSP activities targeted at least one specific criminogenic need, the dichotomous variable was coded as being present. The 5-point variable was coded based on the percentage of needs targeted of the total problematic needs identified on the YCRNA as well as the appropriate prioritizing of need activities. In other words, activities targeting needs that were not rated as problematic or were rated as an area of strength for youth contributed to a lower rating. The variable ranged from no contact with youth to activities addressing at least 50% of youths' problematic needs without addressing any non-problematic needs. The $ICC_{(2,1)}$ value for the variable exceeded the cut-off for excellent reliability at 0.89.

For the balance between support and supervision activities, the variable was coded based on the relative number of activities targeting criminogenic needs, supervision, and non-criminogenic need or other activities. Examples of supervision, non-criminogenic and other activities are included in the coding manual. Where the number of activities of each was comparable or the number of need activities exceeded those of supervision and other activities, the dichotomous variable was rated as present. The 5-point variable ranged from no contact with the youth to the number of criminogenic need activities exceeding the total of supervision and other activities. The $ICC_{(2,1)}$ value for the variable was 0.81, which exceeds the cut-off for excellent reliability.

Results

Descriptive Data. A mean of 14 (*SD* = 17.93) specific activities were documented for youth during ISSP, with up to 120 separate activities identified in a file. An average of 5.45 (*SD* = 11.16) criminogenic need activities, 3.77 (*SD* = 4.04) supervision-related activities, and 4.61 (*SD* = 7.72) non-criminogenic or other activities were documented for youth during the program.

Criminogenic needs that were most frequently targeted with program activities were educational and recreational needs for 60.0% and 29.7% of youth, respectively. The needs most frequently not targeted were in the areas of peers (99.3% of youth rated as having the need) and family relationships (87.4% of youth rated as having the need). As with over-identified goals on the referral, criminogenic needs that were not rated as problematic or rated as strengths on the YCRNA were targeted during the program. This was particularly true of educational needs (9.7% of youth) and recreational needs (8.6% of youth).

In terms of matching needs with activities, a mean of 21.9% (*SD* = 22.47) of youths' needs identified as moderately or very problematic on the YCRNA were targeted by program activities. A mean of 23.4% (*SD* = 25.01) of the activities documented pertained to criminogenic needs rated as problematic. For 28.4% of youth, there was no documentation of any activities pertaining to criminogenic needs.

Implementation. Activities targeting at least four criminogenic needs were documented for 18.8% of youth, while 72.0% of youth participated in at least one activity targeted at a criminogenic need. At least one activity pertaining to youths' *specific* criminogenic need areas as identified by the YCRNA was documented for 64.8% of youth, while the remaining 35.2% of programs did not target any criminogenic needs or did not target criminogenic needs identified as problematic for the youth. A balance between support and supervision/other activities was noted in 40.9% of cases.

Outcomes. Logistic regression analyses were used to examine the relationship between program activities and re-offending in the year following ISSP including the covariates identified in the methods section. None of the activity variables were correlated with any of the covariates, such that interaction terms were not needed.

None of the program activity variables were significantly associated with recidivism, although the general criminogenic needs variable approached significance (B = -0.28, SE(B) = 0.17, Wald $\chi^2(1)$ = 2.65, Exp(B) = 0.76, p = 0.10). The relationships were all in the anticipated negative direction with recidivism (see Tables 59-61).

Discussion

The best practice variable for addressing multiple criminogenic needs was present in only 18% of cases, which was to be anticipated since there are no specific guidelines pertaining to the number of criminogenic needs to be targeted by ISSP. A majority of youth participated in at least one activity pertaining to their specific criminogenic needs, while approximately 40% of youth had a balanced focus between criminogenic need and other activities. The finding for the balance of activities is comparable to other evaluations that have found that many program activities involve "feel good" undirected counselling, such as one-to-one discussion and personal advice (Lane et al., 2005). On average, the number of criminogenic need activities slightly exceeded each of the supervision and non-criminogenic need activities. It should be noted, however, that there was significant variability about the mean of the criminogenic need activities and the mean was likely to have been impacted by outliers (e.g., ISSP workers who visited youth at school programs several times a week). Nevertheless, despite the lack of specific direction from the ISSP guidelines, these best practices were reasonably well implemented.

There was considerable variability in the degree to which different criminogenic needs were addressed. Needs such as education were frequently addressed whether problematic or not, as were recreational needs. Other need areas were very infrequently noted as the target of activities, such as peers and family relationships. However, it may be difficult to identify activities to target negative peers and activities may take the form of informal advice-giving that was not included in the coding for the current study, or may not have been documented. Less than one-quarter of youths' identified needs were targeted by any program activities on average.

None of the three program activity variables significantly predicted recidivism in the year following ISSP, although the relationships with offending were in the direction of reducing recidivism and the relationship between recidivism and the criminogenic need

activities variable approached significance. The finding of a weak relationship was surprising given the strength of the need principle in meta-analytic research (e.g., Andrews & Bonta, 2010a). However, given that programs that are based on mentoring relationships also tend to yield positive outcomes (Tolan, Henry, Schoeny, Bass, Lovegrove, & Nichols, 2013), due perhaps to the informal targeting of criminogenic needs, it is possible that the effects of these variables may have been attenuated in the current study due to the overall effect of having a supportive adult ISSP worker.

The current chapter was limited by a lack of documentation of activities that is likely to have contributed to reductions in the effect size for the program activity variables. Oftentimes specific activities were not documented, although it was clear that there had been contact. Also, many of the dynamic need activities identified by Andrews and colleagues (2006), for example, developing problem-solving and self-management skills or generating pro-social alternatives in risky situations, may occur in a more informal fashion during the program that may not have been documented.

Another potential limitation of the current study is that a single YCRNA time point, potentially well before the actual start of ISSP, was used to determine whether activities targeted youths' specific criminogenic needs. Given that criminogenic needs are dynamic (Herrenkohl et al, 2000; van der Put et al., 2012) and some youth participated in the program for over two years, a mismatch with the risk tool may not necessarily have reflected a failure to target youths' specific needs.

Recommendations. The wide variation among the types and composition of program activities for youth suggests a need to provide greater structure and training for ISSP workers. While ISSP should remain flexible, meta-analytic findings indicate that structured programs yield greater reductions in recidivism as compared to unstructured programs (Lipsey, 1992). In particular, it would be beneficial to implement the best practices noted in this chapter and to provide guidance to ISSP workers regarding the types and optimal balance of support, supervision, and other activities in the program. ISSP workers must select the proper interventions to effectively target youths' needs, which may be a challenge if youth need multiple programs or if programs are not available. Risk assessment can help to guide need targets and ISSP workers should aim to comprehensively address needs while prioritizing those that are most problematic for youth.

As was suggested by the current study, it may be more difficult to target some types of criminogenic needs than others. It may be helpful to include training on teaching problem-solving, for example, and to include examples of types of targets from Andrews and Bonta (2010a) in an ISSP manual. It may be useful for training to include information about possible dynamic need targets so that ISSP workers are more able to identify and document what they are likely already doing or, if not, to be able to target these needs in their interactions with youth. This chapter also underscores the need for a community program manual that is frequently updated to help ISSP workers identify appropriate community providers and activities to help target needs. It could be helpful for the ISSP manual activities to be organized according to criminogenic need areas.

Chapter 8.

Best Practice Implementation and Outcomes

An overarching best practice for Youth Justice ISSP, and the focus of the current study, is the importance of proper program implementation. Even an established treatment program may be rendered ineffective if it is improperly implemented (Sherman et al., 1997; Altschulter, Armstrong, & MacKenzie 1999), but proper implementation is still more important in a program for which there are equivocal findings or for programs that are being evaluated. Regardless of the fact that the Youth Justice ISSP guidelines already reflect a number of best practices or may incorporate these in the future, if the best practices are not followed in the actual delivery of the program, there will be no benefit yielded by including them.

Empirical attention to program implementation has tended to lag behind research on treatment needs and strategies, to the point that Gendreau and colleagues (1999) referred to implementation as a forgotten issue in justice intervention. While problems with implementation in youth justice programming are often identified through retrospective recommendations and anecdotes in discussion sections, a number of researchers have attempted to engage in more systematic study of the effects of implementation on program success (e.g., Lowenkamp, Makarios, Latessa, Lemke, & Smith, 2010; Mihalic & Irwin, 2003).

More recently, sufficient numbers of evaluation studies have included implementation data that meta-analysis of these factors has been possible. In the most comprehensive meta-analysis of correctional programming integrity to date, Andrews and Dowden (2005) identified 10 general indicators of well-implemented programs from theoretical and empirical research literature (see also Gendreau et al., 1999). These factors include staff characteristics, such as hiring staff with appropriate interpersonal skills for intervention (e.g., warmth, enthusiasm), as well as training staff well and

providing clinical supervision (Gendreau et al., 1999). Other indicators pertain to the program's characteristics, such as having a specific model of criminal behaviour guiding practice, ensuring adequate dosage, having an evaluator involved in the design, delivery, or supervision of the program, monitoring program process (i.e., implementation), and employing a training manual (Hill, Andrews, & Hoge, 1991; Lipsey, 1995). Still other factors have been statistically associated with more successful programs in the research literature but may not be tenable in practice, such as ensuring programs are new or delivered to small sample sizes (Hill et al., 1991; Lipsey, 1995). However, these variables may actually be proxies for other features of the program as new programs reduce the risk of program drift where changes are made in the program's delivery and small samples allow for greater consistency of the program and closer oversight. From this perspective, there may be other ways to obtain similar results for these variables (e.g., holding regular training sessions, engaging in ongoing evaluation, greater staffing, etc.). Although many of the implementation variables cannot be examined through a file review, the general indicators are able to inform recommendations for ISSP and provide a sense of the scope of the considerations in effective program implementation.

A number of other meta-analyses have examined the role of implementation quality in youth and adult justice programs (Landenberger & Lipsey, 2005; Lipsey, 1995; Lipsey, 1999; Lipsey, 2009; Lipsey & Wilson, 1998; Lowenkamp, Latessa, & Smith, 2006). Across these and the Andrews and Dowden (2005) studies, significant effect sizes were found for implementation variables and, in some cases, the effect sizes were comparable to or exceeded those associated with different types of intervention (Landenberger & Lipsey, 2005; Lipsey, 2009). Of the specific indicators of implementation included in the Andrews and Dowden (2005) meta-analysis, all but dosage and monitoring process were significantly associated with the treatment effect size. However, when all of the implementation variables were included in the analysis simultaneously, only staff selected for their interpersonal skills, the involvement of an evaluator, and small sample sizes remained significant.

In the case of similar programs to ISSP, implementation quality has been found to significantly moderate re-entry program outcomes, such that poorly implemented programs were even associated with increased recidivism (James et al., 2013). Despite

the clear importance of proper implementation, a recurring finding from ISSP-type evaluation studies was that the programs suffered from problems with implementation, particularly in that control groups received comparable levels of service (e.g., Lane, Turner, Fain, & Sehgal, 2007; Lattimore & Visher, 2010; Wiebush et al., 2005). Wiebush and colleagues (2005) empirically examined the relationship of only one of the program's implementation components with recidivism, but where youth at one site received a higher level of service relative to other intensive aftercare youth and controls, their recidivism outcomes were significantly improved.

Few individual studies have focused on the relationship between multiple implementation variables and recidivism outcomes. An evaluation of 58 adult intensive supervision programs operating in the United States collected staff ratings of 15 items pertaining to effective principles of justice intervention (Lowenkamp, Flores, Holsinger, Makarios, & Latessa, 2010). Programs rated as adhering to 40% or more of the items were associated with a significantly decreased likelihood of re-offence.

Although the effect of implementation is well-established in meta-analytic research for ISSP-type and general youth justice programming, studying the implementation of Youth Justice ISSP is important for several reasons. First, evaluation studies often fail to collect implementation data so there is a lack of information on which to conduct further analyses if the program is not successful or for the purpose of meta-analysis. In the Andrews and Dowden (2005) meta-analysis, information for coding each of the implementation factors was available in only 5-59% of studies, while the average number of indicators noted in studies was 3.46 out of a possible 10. For the meta-analysis of re-entry programs, nearly 30% of studies did not provide any information about implementation (James et al., 2013). Interestingly, studies that did not mention implementation had a similar average effect size to those that were well-implemented, although not surprisingly there was much greater heterogeneity in the effect size.

A second reason to study implementation within ISSP is that the implementation findings from meta-analysis have generally been based on questionable data. If any information about implementation is included in studies at all, meta-analyses are often forced to rely on crude indicators, such as approximate completion rates or the presence of any data on implementation (Lipsey & Cullen, 2007). As an example of implementation indicators used in ISSP-type studies, the meta-analysis of intensive re-

entry programs used a three-point index of whether studies reported that the program was well-implemented, had difficulties with implementation, or did not mention implementation quality at all (James et al., 2013).

Even the effect sizes yielded from Andrew and Dowden's (2005) detailed analysis of implementation factors were often coded by simple indications that the elements were present or absent (Lipsey & Cullen, 2007). The implementation indicators may thus be more of a reflection of whether the element was part of the program philosophy and may not reflect the degree to which elements were actually used in practice (e.g., fidelity to the manual for specific program providers, actual dosage). There is a dearth of information on the within-program variability of implementation factors for individuals and there are few studies in which implementation factors are linked to recidivism using actual program data (e.g., Bouffard & Bergseth, 2008).

A third reason to study implementation in the context of Youth Justice ISSP is that the impact of program implementation may be particularly visible in a program like ISSP, which is relatively unstructured outside of basic guidelines and thus allows for considerable variation in delivery. It is likely that implementation quality significantly contributes to the small or null results for ISSP-type programs. Finally, a study of ISSP implementation can provide valuable feedback for the program and support ISSP workers in attending to implementation quality in their program delivery.

Another consideration for Youth Justice ISSP is whether there may be specific characteristics that enable youth to benefit more or less from the program. Under the RNR model (Andrews et al., 1990), specific responsivity refers to matching services to youths' motivation level, learning styles, or demographic factors. While some responsivity factors could be addressed in the way Youth Justice ISSP is delivered, demographic factors could inform which youth are selected for the program and those who might be better diverted to other resources. Unfortunately, specific responsivity factors have yet to be explored in the general meta-analytic literature or to be subject to replication (Andrews, Bonta, & Wormith, 2006). In the James and colleagues (2013) meta-analysis of re-entry programs, effect sizes were found to be larger for older youth, violent youth, gang-involved youth, and for populations with a larger proportion of males and youth from an ethnic minority group. Effect sizes were smaller for studies with a

higher proportion of youth who abused substances. It is possible to examine some of these factors in the current study to determine whether implementation might be particularly important for certain groups.

A final implementation consideration is the importance of youths' positive engagement with ISSP. In speaking with Youth Justice ISSP staff during a prior program review (Mordell, van der Woerd, Viljoen, & Roesch, 2008), emphasis was often placed on the importance of the relationship with youth. In addition to having support to attend appointments and opportunities for informal targeting of criminogenic needs, a stable, warm relationship with a caring adult may act as a protective factor for youth violence (Simons, Paternite, & Shore, 2001).

Evaluating program design and delivery. Since proper implementation is an implicit assumption of programs, ISSP guidelines pertaining to proper implementation do not exist. However, in that ISSP workers are meant to document their contacts, it appears that Youth Justice ISSP has an interest in monitoring program process. Unfortunately, given the quality of the documentation on the file, there does not appear to be much follow up for the program's implementation. Youth Justice ISSP does not currently evaluate either the program process or outcomes and adherence data are not formally or routinely collected. Best practice research has been incorporated into a number of the program guidelines, although there are still best practices that have not been adopted.

For Youth Justice ISSP, there are a few ways that the study of implementation may be approached. The first of these is to examine program delivery according to the ISSP guidelines, which forms the basis of the final best practice variable. A second means of studying implementation is according to a best practice framework of program delivery, for which a best practice composite was calculated. A third measure was created to only contain variables pertinent to program experience. This composite was created because it was missing fewer data and because the referral variables are more removed from program experience. Thus, a youth who did not appear to receive any program at all could still receive a score of four for the complete best practice composite. The program-only variable was calculated for comparison purposes with the full variable, to determine whether it was better able to predict outcomes.

For the implementation of the program, a reasonably high level of adherence was anticipated for program guidelines since these are included in an ISSP manual and ISSP workers receive annual training. Since ISSP workers may not be familiar with best practice research for ISSP-type programs, a lower level of fidelity to best practices was expected. Based on Lowenkamp and colleagues' (2010) findings for integrity to effective intervention principles, fidelity might be expected for approximately one-quarter of the best practices not included in the program guidelines.

It was of interest to determine whether ISSP implementation was related to a number of both positive and negative non-recidivism outcomes. It has been suggested that broad recidivism outcomes alone are often insufficient to determine whether an intensive support program is making an impact (Wiebush et al., 2005). Better implementation was anticipated to be associated with positive changes in intermediate targets in the relationship with recidivism, such as community program participation, employment, and school outcomes. Other probation outcomes such as breaches, police contact, and unofficial violence and substance use were expected to be negatively related to ISSP implementation.

One of the key research questions in the current study was whether youth receiving a better-implemented ISSP according to best practices were less likely to re-offend. While a significant relationship would provide evidence for the importance of proper implementation, it could also provide an indication of the effectiveness of ISSP since even a well-implemented ineffective program would not yield any effect for recidivism. Since implementation was central to the current study, multiple indices of recidivism were included to provide a more sensitive measure of the program's effect, as it may be unrealistic to expect a program to prevent any recidivism for youth at high risk of re-offending. These measures included number of convictions, types of offences, incidences of offending, and an index of seriousness. As ISSP may confer longer-term benefits for reductions in recidivism through its effects on intermediate targets, measures of any recidivism post-follow up and degree of justice system involvement at the time of recidivism data collection were also used.

The current study sought to identify potential responsivity factors for the relationship between implementation and any recidivism, such as gender, age, Aboriginal identification, and substance abuse to provide preliminary indications of what

factors may be of interest to study prospectively in a Youth Justice ISSP evaluation. Meta-analytic research has found age, ethnicity, violence, and addictions to moderate effect sizes for intensive re-entry programs (James et al., 2013). Limited research has been conducted on the effect of gender on ISSP and many evaluation projects have been limited to males. A final variable of interest pertaining to recidivism was youths' engagement with the ISSP worker.

Procedures

The coding for the dichotomous Youth Justice ISSP guideline adherence variable was based on the presence of all seven program guideline variables. These were:

a) Select youth according to program criteria*
b) Target individual and systemic factors
c) Make contact with ongoing or new community resources
d) Involve youths' families
e) Duration of a minimum of six months*
f) Intensity of eight hours per month*
g) Activities target criminogenic needs

The 5-point variable was coded based on the total number of guidelines to which the program adhered. The $ICC_{(2,1)}$ value for the Youth Justice implementation variable was 0.85, which exceeded the cut-off for excellent reliability.

Best practice fidelity was measured using the total of the dichotomous ratings for the best practice variables in the previous chapters, as well as the ISSP adherence variable above. In cases where ISSP guidelines and best practices differed, marked with an asterisk above and below, the best practice ratings were substituted in the composite variable.

In addition to the best practices captured by the adherence model above (those not marked with an asterisk), the complete fidelity model comprised:

a) Select high-risk youth for the program*
b) Use a systematic risk assessment tool to identify youth
c) Use systematic risk assessment to identify program goals
d) Make youths' transition to school or the workforce a priority

e) Intensity is a minimum of 100 hours*
f) Duration of nine months*
g) Target criminogenic needs specific to each youth
h) Ensure a balance between support and supervision activities
i) Program delivery adheres to guidelines

The $ICC_{(2,1)}$ value for the best practice composite exceeded the cut-off for excellent reliability at 0.92.

Responsivity factors were coded through database and file information. Where pre-existing groups did not exist, they were created through the use of a cut-point at the 50^{th} percentile of values (e.g., for age, SAVRY risk score). It has been suggested that a minimum sample size of 100 is required for logistic regression (Long, 1997). Since many of the variables of interest fell well short of this, only those with empirical support and the main demographic variables (gender, age) were run with the caution that there may be insufficient power to detect an effect. Although ethnicity has been found to moderate offence outcomes in the past (James et al., 2013), an analysis of all of the ethnic backgrounds represented in the study was not feasible. The focus of this responsivity factor was on whether or not youth were identified as Aboriginal, given the overrepresentation of Aboriginal youth in the criminal justice system and the implications for service delivery. Variables for which the sample size in one of the groups was considered insufficient to run the analyses were for youth documented as having FASD, low cognitive functioning, mental health issues, and youth with no history of violence.

Youth's engagement with ISSP was determined through a crude rating of whether YPOs described youths' relationship with ISSP in positive, neutral, or negative terms. For example, probation notes may have indicated "youth is sad that ISSP is ending" (which would be coded as positive), or "youth does not get along with ISSP" (which would be coded as negative). Probation notes without any valence were coded as neutral. See Table 4 for the inter-rater reliability data for the engagement variable.

Results

Implementation. The average adherence score (out of seven) for the Youth Justice ISSP guidelines was 4.10 (*SD* = 1.86) or 58.5%. The program appeared to have

been fully implemented according to guidelines for 10.9% of youth, while 14.3% of files showed evidence of adhering to only one or no ISSP guidelines.

For fidelity to a best practice model, the average implementation score out of 13 was 6.64 (*SD* = 2.82) for the 140 youth for which an implementation score could be calculated (i.e., file was not missing referral or YCRNA) or 51.2%. Fidelity to the complete best practice model was only recorded for 2 youth (1.8%) in the program. For the alternate, program-only implementation composite, which was missing fewer data, the mean score was 5.01 (*SD* = 2.79) out of a possible 11, while 3 (1.7%) of 170 youth received a program consistent with the program-only best practice model

Outcomes – ISSP guideline adherence composite. In the interest of clarity, adherence to the ISSP guidelines was treated as the other program variables in the study, where analysis was limited to the relationship with dichotomous recidivism. Hierarchical logistic regression analysis was used to examine the relationship between the 5-point ISSP adherence variable and re-offending in the year following ISSP. The adherence variable was not significantly associated with recidivism (see Table 62).

Intermediate outcomes – Best practice fidelity composite. Hierarchical logistic and linear regression analyses were used for dichotomous and continuous outcome variables, respectively. Interaction terms were created for the implementation composite and file outcome quality variable ($r = 0.25$, $p < 0.01$) and the file time variable ($r = 0.58$, $p < 0.001$) for use in moderation analyses for the relevant models. An interaction term was also created for the implementation composite and age ($r = -0.36$, $p < 0.001$), although the variables were first centered due to the lack of a meaningful zero point for age and to avoid problems with multicollinearity.

In terms of employment outcomes, implementation was marginally associated with obtaining part-time employment while on ISSP or during the follow-up year ($B = 0.15$, SE(B) = 0.09, Wald $\chi^2(1) = 2.76$, Exp(B) = 1.16, $p = 0.10$) prior to the addition of the file time interaction term to the model, which was not significant. Any full-time employment and overall employment during ISSP and follow up, as well as employment at the end of ISSP were not significantly related to implementation. See Tables 63 to 66 for the complete models for employment outcomes.

With respect to school outcomes, the best practice composite showed a marginally significant association with overall enrollment in school during ISSP and follow up prior to the addition of the interaction terms to the model (β = 0.16, SE(B) = 0.02, t(123) = 1.81, p = 0.07). The interaction term for age and best practice implementation was not significant. Implementation according to best practices was not associated with any school enrollment during ISSP or follow up or with enrollment in school at the end of the follow-up period. See Tables 67 to 69 for analyses of the school outcomes.

Regarding participation in community programs, the relationship between implementation and the number of programs that youth received on ISSP and in the follow-up year was not significant (see Table 70).

Probation outcomes – Best practice fidelity composite. Hierarchical logistic and linear regression analyses were used for dichotomous and continuous outcome variables, respectively, for multiple probation outcomes documented in the file. Covariates for each of the probation outcomes were identified through correlation (see Table 71) and are included in the tables of the analyses.

Implementation quality approached significance in its relationship with failures to report to probation appointments (β = 0.15, SE(B) = 0.02, t(127) = 1.71, p = 0.09). The number of breaches of conditions detected by probation officers, curfew violations, and periods of unknown whereabouts were not associated with the implementation composite. See Tables 72 to 75 for the complete analyses.

Files of youth receiving a better implemented program were marginally more likely to have documented incidences of violence (β = 0.14, SE(B) = 0.02, t(121) = 1.67, p = 0.10), although the relationship was no longer significant once the age and file quality interaction terms were added to the model. Implementation was also significantly positively related to documented incidences of substance use (β = 0.16, SE(B) = 0.02, t(125) = 2.05, $p < 0.05$), even with the file quality interaction added to the model. The best practice composite was not significantly related to file documented police contact. See Tables 76 to 78 for the complete models for violence and substance use.

Recidivism outcomes – Best practice fidelity composite. Hierarchical linear regression was used for the analysis of the implementation composite, aside from

dichotomous recidivism for which hierarchical logistic regression was employed. Covariates with the recidivism variables were identified through correlation analyses (see Table 6) and are identified in the tables of the models. Although the SAVRY was not significantly related to each of the recidivism variables, it was nonetheless included in all of the analyses due to the importance of controlling for risk. An index of offending of the same type (e.g., violent, non-breach, etc.) in the year prior to ISSP was also included as a covariate for the variables where available.

Implementation according to a best practice model of ISSP was not significantly related to dichotomous recidivism (see Table 79). To test whether the program-only composite would provide greater explanatory power in the relationship with recidivism, it was run in a separate analysis. However, the results suggested that the program-only composite was roughly equivalent in its ability to predict recidivism (see Table 80). Thus the full program implementation variable was used for the remainder of the recidivism analyses since it is most reflective of a complete model of best practice.

To further examine potentially subtler effects of ISSP implementation, other measures of recidivism in the year following ISSP were subjected to analysis. The best practice fidelity composite was negatively associated with a total seriousness index in the year following ISSP (β = -0.17, SE(B) = 0.03, t(123) = -2.06, p < 0.05). Implementation also marginally predicted the number of non-breach convictions (β = -0.16, SE(B) = 0.04, t(123) = -1.79, p = 0.08) and time out of custody (β = 0.15, SE(B) = 2.53, t(129) = 1.82, p = 0.07). Implementation did not predict the number of convictions, the number of offence dates, the number of violent convictions or indictable convictions, although all relationships were in the anticipated negative direction. See Tables 81 to 87 for the complete models.

With respect to the long-term effect of implementation on recidivism outcomes, the post follow-up period ranged from zero to 1400 days. Youth in custody for the entire period post-follow up or who had a follow-up period of zero days were excluded from the analyses given the lack of opportunity to reoffend (n = 3). Implementation quality was not associated with any recidivism post follow up or the degree of current justice system involvement (see Tables 88 & 89).

Responsivity factors for best practice fidelity composite. In order to explore whether there may be certain responsivity factors for ISSP implementation, group

analyses were conducted for recidivism in the follow-up year. Hierarchical logistic regression was employed with only the SAVRY total score entered in the first block, since the analyses were already underpowered due to the smaller sample sizes. Groups were run separately to determine if there was an effect of implementation for recidivism with the sample sizes for each noted below.

There was no significant effect of implementation on recidivism for either male or female youth when the groups were run separately (see Table 90). The dichotomous variable created for age with the cut-point at the 50th percentile corresponded to approximately 16 years and 4 months. There was a significant negative association with recidivism for implementation for the younger aged participants (n = 66; B = -0.24, SE(B) = 0.12, Wald $\chi^2(1) = 4.45$, Exp(B) = 0.78, $p < 0.05$) but not the older aged group (see Table 91). Better implementation quality predicted decreased reoffending for Aboriginal youth (n = 32; B = -0.46, SE(B) = 0.19, Wald $\chi^2(1) = 5.69$, Exp(B) = 0.63, $p < 0.05$), while implementation was not significantly associated with recidivism for youth with no or unknown Aboriginal identification (see Table 92).

In terms of other factors that might act as responsivity factors for program implementation, risk was dichotomized using the SAVRY value falling at the 50th percentile (32) as the cut point. High-risk youth showed a significant effect for implementation (n = 64; B = -0.25, SE(B) = 0.11, Wald $\chi^2(1) = 5.55$, Exp(B) = 0.78, $p <$ 0.05) while low-risk youth did not (see Table 93). There was no significant effect of implementation whether youth were coded as having substance use problems or not (see Table 94).

Engagement. A final factor pertaining to implementation is the extent to which youth appear to have engaged in the program. While many factors could contribute to engagement, this variable provides a measure of the effect of the relationship between youth and ISSP workers on recidivism. Engagement in ISSP was significantly associated with reductions in recidivism in the following year (B = -0.56, SE = 0.28, Wald $\chi^2(1) = 4.18$, Exp(B) = 0.57, $p < 0.05$). See Table 95 for the complete model.

Discussion

The current study found that, on average, individuals received a program adhering to 58% of the Youth Justice ISSP guidelines. Only approximately 11% of participants received the program as it was fully intended to be delivered. A similar degree of fidelity to best practices was found, although it should be noted that a number of the best practices overlapped with Youth Justice ISSP guidelines. Focusing on the best practice items that were not stipulated in program guidelines but constituted more general principles of effective intervention as in Lowenkamp and colleagues (2010), the average level of fidelity to best practices was still in the range of 50%.

Better implementation was associated with intermediate outcomes in the relationship with recidivism, including obtaining part-time employment and overall school enrollment. However, proper implementation was also positively associated with a number of probation outcomes that were not anticipated, such as drug and alcohol use. It has been suggested that ISSP-type programs may lead to higher numbers of breach charges due to the increased level of supervision (Altschuler & Armstrong, 1999). Similarly, it is possible that the higher level of contact associated with better-implemented programs led to the greater detection and subsequent documentation of these outcomes. It is also possible that data quality issues contributed to the shared variance for these variables, although attempts were made to control for these.

Fortunately, recidivism outcomes were not reliant on file quality and were thus able to provide more reliable data. A total index of seriousness in the year following ISSP was significantly associated with implementation, while other recidivism outcomes were marginally associated with the best practice composite (number of non-breach convictions and days out of custody during the follow-up year). For the recidivism outcomes that were not significantly associated with ISSP implementation, the relationship was still negative, with the effects of the program appearing to attenuate over time. Other evaluations of ISSP-type programs with short- and longer-term follow-up periods have similarly found that the effects of programs diminish over time (Abrams et al., 2011; Drake & Barnoski, 2006). Given that the ISSP documentation was very limited, it is nonetheless encouraging that a trend towards the reduction of recidivism persisted across measures. It must also be considered that effect sizes for juvenile justice programs are often small, such that significant effects might only appear in meta-

analyses with samples sizes of thousands of youth (e.g., James et al., 2013; Lipsey, 1999). Moreover, the negative relationship with recidivism for implementation was found within program participants, such that even participants receiving a low implementation rating may have actually received at least some programming. Studies in which the control group appeared to have received comparable levels of services to the program group have similarly failed to yield significant findings for recidivism (Lane et al., 2007).

A number of potential responsivity factors for implementation were identified in the current study, although it should be noted that significant findings indicate that implementation may be particularly critical among certain groups and the ability to extrapolate the findings to suggest that the program was more effective for these groups would be extremely limited. Also, prior to using the findings in any decision-making capacity, it is important that the results be confirmed through prospective research. The results suggest that, while proper implementation should be a focus for all youth, particular attention might be paid to the implementation of the program for younger, for Aboriginal, and for higher-risk participants.

Given the potential significance of the finding of a unique effect of implementation for Aboriginal youth, an ad-hoc Aboriginal advisory committee was formed for consultation in interpreting the findings. Several possible mechanisms were proposed by the committee, including the possibility that more social supports, better family support, employment, or increased access to cultural programming may be associated with decreased recidivism, either causally due to ISSP or as correlates of better implementation that mediate the relationship with recidivism. Unfortunately, for the most part, it was not possible to test these hypotheses based on the available data. It should be noted that identification as Aboriginal does not provide any information about access to resources, cultural identification, etc., such that it is difficult to make conclusions about the findings.

Since Aboriginal identification was significantly associated with crude indicators of FASD, low cognitive functioning, and addictions, these variables were added to the recidivism model for Aboriginal youth, but implementation remained significant. The finding of an effect for implementation for Aboriginal youth is consistent with other research indicating that ISSP-type programs show a greater effect size for youth in marginalized groups (James et al., 2013). It is possible that youth not identified as

Aboriginal have greater access to resources regardless of their participation in ISSP, such that the quality of the program delivery may not have as significant an impact on outcomes. However, for Aboriginal youth, who may otherwise have fewer opportunities to access such resources, ensuring that program delivery is consistent with best practice guidelines may be particularly crucial. When analyses of the contextual best practice variables were run separately for Aboriginal youth, these variables significantly predicted recidivism in the year following ISSP; however, the same was true of nearly all of the best practice variables. Thus, while an increase in community supports may contribute to the relationship between implementation and recidivism for Aboriginal youth, it does not appear to account for the whole relationship. Further prospective study is needed to explore factors that could not be reliably or consistently coded from the files.

As to a finding for implementation for younger participants, it may be that ISSP operates through different mechanisms depending on age (e.g., mentoring function more important for older participants, community engagement function more important for younger participants) or that other factors are influencing the relationship between implementation and recidivism for younger and older participants. In Chapter 5, it was found that age was negatively correlated with three of the contextual variables, such that older participants may have had a more restricted range with respect to these variables. The finding of a significant effect for higher-risk youth is consistent with the RNR model (Andrews et al., 1990) and would appear to confirm that these youth have a higher level of needs that are more likely to be impacted by proper implementation (Bonta & Andrews, 2007).

Regarding limitations of this chapter, the file review methodology could not rule out potential third variables, such as youth factors that may make them more amenable to participating in ISSP and thus to receiving a better-implemented program as well as being successful in other outcomes. In particular, the engagement variable may actually reflect youth characteristics that are responsible for the decrease in recidivism. Without more information about youth characteristics or a control group, the conclusions that can be drawn from the findings are limited. Furthermore, much of the information necessary for making best practice ratings may not have been documented in the file, such that fidelity may have been underestimated in a large number of cases. The fact that there

were still significant findings for implementation despite the numerous problems with the data suggests a robust effect for implementation.

Also, the current study was limited in its ability to examine responsivity factors for implementation, both in having insufficient sample sizes for some variables of interest that were included in the study and from variables that could not be extracted from files, such as motivation. The results are not able to provide a measure of responsivity for the program in general, although significant findings associated with implementation may provide some measure of the effectiveness of the program for particular groups.

Recommendations. Findings from the study indicate the importance of both proper implementation and youth engagement. The latter is in line with recent changes being made in Youth Justice to implement a motivational enhancement protocol for YPOs. Incorporating aspects of motivational interviewing could serve to increase the engagement of youth in the program to both facilitate the better delivery of the program as well as provide a supportive relationship for youth.

To support proper implementation, Youth Justice ISSP would benefit from collecting process data to encourage greater adherence to guidelines within the program as well as having the data available for evaluation purposes. While Youth Justice ISSP has already adopted many of the best practices discussed in the current study as part of their program guidelines, the effect size for implementation suggests that incorporating all of the best practices into the program guidelines may lead to improved recidivism and other outcomes. ISSP workers may benefit from additional training in best practices and well as support in implementing them (e.g., development of an ISSP manual).

It is also important that potential responsivity factors for the implementation of ISSP are studied further. While fidelity to a best practice model of implementation should be the goal for all participants, it may be particularly important for younger participants, youth identified as Aboriginal, and high-risk youth. Though the study findings may not suggest any program changes, the finding of a significant effect for implementation for Aboriginal youth suggests a need to seek consultation from Aboriginal stakeholders as to possible means of improving program practice. As an example, an Australian intensive supervision program includes an Aboriginal team advisor who engages and works with Aboriginal families because of the high percentage of Aboriginal youth in the program (Department of Corrective Services, 2007).

Chapter 9.

Discussion

In an effort to reduce the reliance on custodial measures for adolescents, the *Youth Criminal Justice Act* (2002) designated the Intensive Support and Supervision Program (ISSP) as an alternative to custody or a means to shorten custody sentences and support re-entry to the community (Sec. 41(2)(l)). Youth Justice's ISSP is strong in its guiding principles and program delivery guidelines, incorporating a number of best practices. Other best practice research has emerged from empirical and review studies of similar programs to Youth Justice ISSP and has become more widely available since the program's inception that does not appear to be reflected in the program's current design. In particular, proper implementation is not a strong focus of ISSP and the way the program is offered to youth may be quite different than how it appears on paper.

In terms of implementation, adherence to an average of nearly 60% of program guidelines was reflected for individual youth in practice. Only a small percentage of youth were considered to have received a fully-implemented ISSP according to the program guidelines. Since ISSP workers may not be familiar with evidence-based best practices for ISSP-type programs, the fidelity to approximately 50% of best practices in the program's delivery was higher than anticipated. Though research for individual-level best practice fidelity is limited, this figure compares quite favourably to the findings of 25% fidelity to principles of effective intervention by programs for the Lowenkamp and colleagues (2010) study.

Although none of the individual best practice variables were significantly associated with reductions in recidivism, nearly all showed a negative relationship with recidivism in the year following ISSP. Additionally, actual program duration in the community predicted reductions in recidivism, while the number of hours of ISSP and the criminogenic need activities variable were marginally related to recidivism. A number of

contextual variables were associated with intermediate targets such as part-time employment and school enrollment while on ISSP or in the year following ISSP.

A composite of all of the best practices was significantly associated with decreased total seriousness of offences in the year following ISSP and marginally associated with increased time in the community and decreases in the number of non-breach convictions. The best practice fidelity composite was also associated with intermediate employment and school outcomes. While not as conclusive as a control group evaluation, the current study does support that ISSP may be effective if properly implemented, as suggested by an overall negative effect size of implementation on all recidivism measures in the year following ISSP.

A significant finding from the study was that a properly-implemented ISSP may be particularly beneficial for Aboriginal youth. At the same time, the number of youth identified as having any Aboriginal background in the study comprised only 23% of participants, which is discrepant with estimates of one-third of youth in the youth justice system and 47% of youth in custody being identified as Aboriginal (McCreary Centre Society, 2005; Representative for Children and Youth & Office of the Provincial Health Officer, 2009). While it is unclear how Aboriginal identification was determined in each case, and a significant number of youth in the study were classified as having unknown Aboriginal identification, it is important that the reasons for this discrepancy are explored further. It may be that Aboriginal youth were more likely to be referred to community providers of ISSP. In the probation survey, YPOs identified a preference for community ISSP where there was an Aboriginal ISSP worker, or it may be that youth are referred to other programming. The study results also suggested that proper implementation may be particularly important for younger participants and higher-risk youth.

Limitations of the current study

Procedural disadvantages. Due to the retrospective nature of the study, the conclusions that may be drawn from the findings are limited. It was not possible to include other variables that could influence or even account for the relationship between ISSP and outcomes, including staff and youth factors. There were also some best practices that could not be the subject of a file review, such as ensuring that staff use individualized case planning (Gies, 2003), use rewards and appropriate sanctions

(Goodstein & Sontheimer, 1997), and are well-trained and well-supported (Petersilia & Turner, 1991).

Issues pertaining to the sample also created problems for the study. The sample size, though estimated to be adequately powered for the analyses, may not have been sufficiently large to detect small effect sizes, in particular for the smaller groups used to examine responsivity factors for implementation. The sample size did not allow for any statistical corrections for the multiple analyses conducted, increasing the likelihood of yielding a significant finding by chance. In terms of the youth that had to necessarily be excluded from the sample, those who were excluded through the non-disclosure mechanism may have biased the results of the study. If a well-implemented ISSP did serve to reduce recidivism, youth receiving a well-implemented program may have been disproportionately represented in the youth who successfully completed the non-disclosure period.

While the best practice fidelity composite predicted outcomes, it should be noted that a validated framework for ISSP evaluation does not currently exist. As such, although the choice of variables for inclusion in the index was achieved through a systematic process, the composition of the index itself was not empirically-based as there is a lack of research to inform the specific elements critical to evaluation. For example, there was some redundancy among the four contextual variables that was likely to inflate the contribution of the contextual best practices to the index as compared to others. The SPEP guidelines (Lipsey et al., 2010), for example, are weighted based on the relative contributions of the variables to recidivism in the meta-analytic research. Component analyses for intervention from the behavioural analysis literature may be useful in examining the individual effects of components, the interactive relations among components, the necessary and sufficient aspects of components, the effects of combining components, and the effects of sequencing components in a particular order (Ward-Horner & Sturmey, 2010).

Data quality problems. The validity of the results of the study is tied to the quality of the data from which the results were drawn. The current study encountered problems with missing data, data that were insufficient to address the research questions, and inaccurate data. Efforts were made to minimize the threats to validity that these issues posed including the careful consideration of how missing data issues

would be handled, coding program information and outcomes based only on what was documented, and augmenting data quality through the use of multiple sources. However, due to missing data, approximately one-quarter of cases were omitted from the primarily analyses for implementation, resulting in analyses that had significantly reduced power to detect effects. While pilot imputation procedures and alternate indices with fewer missing data yielded similar estimates, the overall effect of implementation may have been reduced by the high number of missing cases.

It is important to note that data quality issues were pervasive across sources and types of information. An evaluation of ISSP based solely on the database information without conducting a thorough scan of the file data, for example, would be based on flawed assumptions of the program's actual delivery. While some of the inconsistencies may have minor implications for the findings and operation of ISSP, others such as the discrepant results for ISSP time in the community and total ISSP time yielded significant differences for recidivism outcomes. The effects of data quality in the current study have implications for the broader literature in terms of the failure to find effects or the inconsistencies in effect sizes for best practices across studies of ISSP-type programs.

Of the data management strategies selected for the study, the overall result was that information was coded more conservatively than what is likely to have been the case in reality. On the one hand, this strategy may serve to underestimate the occurrences of undesirable file outcomes (e.g., breaches, violence). However, the more significant consequence to this study is that the program captured provides a minimum estimate of the functioning of the actual program and that the effects of the best practices are more likely to be attenuated. While a conservative strategy increases the risk of failing to detect an effect where one exists, this is a preferable outcome to falsely attributing effects to best practices that were not actually present by being overly generous in coding information or imputing values. Nevertheless, the findings of the study should be viewed with the caution that the effects of best practices and the program may be underestimated.

Inter-rater reliability concerns. Two problems complicated the file inter-rater data for the current study. Unfortunately, due to resource limitations and other commitments, it was not possible to have a fully crossed research design to evaluate the reliability coefficients for all variables extracted from the files. This may have led to the

underestimation of reliability coefficients for outcomes due to the inability to model main effects from raters. Although there are options available to correct for the bias in non-fully crossed designs (e.g., Putka, Le, McCloy, & Diaz, 2008), these are not readily available in statistical packages and the calculation of the coefficients in samples with large numbers of variables is unwieldy. As conventional reliability statistics were still able to yield estimates that met the cut-off criteria for the study, these corrections were not deemed necessary.

A second problem was the small samples sizes for the file inter-rater analyses. This was less problematic for the inter-rater agreement analyses, for which the use of multiple variables to calculate estimates was able to generate more observations for the raters. The outcome variables had a maximum of 10 observations per variable, which greatly reduced the precision of these estimates. The use of a small sample size has the effect of creating very large confidence intervals, for which some in the current study included negative values. The potential impact of low inter-rater reliability on the results of the study for these variables should be considered due to the increased likelihood of failing to detect an effect where one exists (Hallgren, 2012).

Core Recommendations

One recommendation that is clearly indicated by the findings of the current study is the need for a prospective evaluation of the program. A prospective evaluation could both collect a wider, more targeted range of data as well as provide greater confidence that data quality or other factors were not diminishing or altering the effect size for implementation. The results of the current study were encouraging with respect to the effectiveness of ISSP and the importance of implementation, but due to the limitations associated with a retrospective file study, it is difficult to conclusively attribute these results to the program. Component analyses, or more specifically parametric analyses that rely on levels of program elements rather than their absence or presence, might be usefully employed to elucidate the necessary and sufficient aspects of ISSP. While a full parametric analysis was outside of the scope of the current study, the process of developing a framework was intended to lay the foundation for prospective evaluation in which the preliminary framework could be further tested and optimized.

Whether or not an evaluation may be supported at this time, a second recommendation is to implement measures to improve ISSP documentation. While the current study relied upon older files, comments from the probation survey would suggest that practices have not improved markedly in this respect. It is not clear whether the problems with documentation are that logs are not completed by ISSP workers or whether completed logs are not placed on the probation file by YPOs, although the former appears to be more likely given that the proportion of logs on file was fairly consistent among specific ISSP workers. In general, ensuring that data are accurate across sources (e.g., database and file) is important since the database is more likely to be used in the reporting of summary data. It is critical that this particular weakness of the program is addressed for audit and accreditation purposes as well as general accountability.

It would also be beneficial for Youth Justice to revise their ISSP guidelines to be informed by the complete framework of best practice research (Lipsey et al., 2010). Where non-guideline best practices were nonetheless implemented by staff, they were associated with improvements, if small, in a variety of outcomes. While the flexibility of ISSP is a key strength of the program, the fine line between flexibility and insufficient structure may become more blurred in practice without clear guidelines and monitoring. Further recommendations pertaining to specific best practices are included in Chapters 3 to 8 above.

A final overall recommendation for ISSP is for a program coordinator to be hired whose sole responsibility would be to oversee ISSP program delivery, including engaging in ongoing evaluation, conducting training, updating community resources and program manuals, and providing clinical support to ISSP workers (Andrews & Bonta, 2010b). Although these activities could help to support ISSP workers deliver an even better program, the duties would be burdensome to staff for which ISSP is an adjunct responsibility. An additional benefit to hiring a program coordinator is that having an evaluator involved with programs improves outcomes (Andrews & Bonta, 2010b).

Future Research

Further research in this area is necessary in order to address the shortcomings of this study. As noted in the previous section, a prospective study of ISSP where youth

are recruited from their enrolment in the program with a non-ISSP control group would allow for conclusions to be drawn about the effectiveness of the program with more certainty. Such a study could focus on recidivism as well as more general criminogenic needs. Other best practices that could not be studied through file review might be studied using an interview or questionnaire study of probation officers and ISSP workers.

In terms of the general research literature, the assumption often seems to have been that difficulties or inconsistencies in implementation are part of the statistical "noise" of a program and that good programs should be robust enough to yield an effect despite these factors. It may be that many existing programs that have returned equivocal results suffered from problems in implementation and that programs could be improved rather than replaced by a model program. More research is needed on the effects of implementation, particularly with respect to specific best practices operating at the level of individual participants. The quality of data used by meta-analyses for implementation and for the best practices included in the current study has tended to be quite poor and findings from meta-analyses may have underestimated the importance of certain variables due to the wide variation of implementation within studies. With the debate about whether treatment works for youth offenders more or less resolved, research is needed to identify what works for whom under which circumstances, for which implementation is a key component.

Conclusion

ISSP is currently being offered to youth at high risk of reoffending as an alternative to custody or a means of facilitating re-entry to the community. However, the program allows for considerable flexibility in its delivery, which can lead to inconsistent program delivery and outcomes. The current study indicated that the program is well-implemented on a number of dimensions and yielded promising findings for a best practice framework of practice for ISSP across a number of outcomes. The study represents an important first step to more rigorous evaluation research to ensure that ISSP is improving the lives of high-risk, justice-involved youth.

References

Abrams, L. S., Terry, D., & Franke, T. D. (2011). Community-based juvenile re-entry services: The effects of service dosage on juvenile and adult recidivism. *Journal of Offender Rehabilitation, 50,* 492-510.

Altschuler, D. M., & Armstrong, T. L. (2004). *Intensive juvenile aftercare reference guide*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention. Retrieved August 19, 2011 from https://www.csus.edu/ssis/cdcps/intensiveaftercarereferenceguide.pdf.

Altschuler, D. M., & Armstrong, T. L. (1999). Reintegrative confinement and intensive aftercare. *Juvenile Justice Bulletin*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention. Retrieved June 7, 2008 from https://www.ojjdp.gov/jjbulletin/9907_3/reintegrate.html.

Altschuler, D. M., & Armstrong, T. L. (1994). *Intensive aftercare for high-risk juveniles: A community care model*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention. Retrieved May 29, 2008 from https://www.ncjrs.gov/pdffiles/juvcc.pdf.

Altschuler, D. M., Armstrong, T. L., & MacKenzie, D. L. (1999). Reintegration, supervised release, and intensive aftercare. *Juvenile Justice Bulletin.* Washington, DC: Office of Juvenile Justice and Delinquency Prevention. Retrieved June 7, 2008 from https://www.ncjrs.gov/pdffiles1/175715.pdf.

Altschuler, D. M., & Brash, R. (2004). Adolescent and teenage offenders confronting the challenges and opportunities of reentry. *Youth Violence and Juvenile Justice, 2*, 72-87.

Andrews, D. A., & Bonta, J. (2010a). *The psychology of criminal conduct (*5th ed). Cincinnati, OH: Anderson Publishing Company.

Andrews, D. A., & Bonta, J. (2010b). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and Law, 16*, 39-55.

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior, 17*, 19-52.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The recent past and near future of risk and/or need assessment. *Crime & Delinquency, 52,* 7-27.

Andrews, D. A., & Dowden, C. (2005). Risk principle of case classification in correctional treatment: A meta-analytic investigation. *International Journal of Offender Therapy and Comparative Criminology, 50,* 88-100.

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychological informed meta-analysis. *Criminology, 28,* 369-404.

Anthony, E. K., Samples, M. D., de Kervor, D. N., Ituarte, S., Lee, C., & Austin, M.J. (2010). Coming back home: The reintegration of formerly incarcerated youth with service implications. *Children and Youth Services Review, 32,* 1271-1277.

Aos, S., Phipps, P., Barnoski, R., and Lieb, R. (2001). *The comparative costs and benefits of programs to reduce crime, Version 4.0.* Olympia, WA: Washington State Institute for Public Policy. Retrieved January 3, 2014 from https://www.wsipp.wa.gov/Reports/01-05-1201.

Australian Human Rights Commission. (2008). *Preventing crime and promoting rights for indigenous young people with cognitive disabilities and mental health issues.* Sydney, NSW: Australian Human Rights Commission. Retrieved August 20, 2011 from https://www.humanrights.gov.au/publications.

Baird, C. (1991). Intensive supervision programs for high-risk juveniles: Critical issues for program evaluation. In T. L. Armstrong (Ed.), *Intensive interventions with high risk youths: Promising approaches in juvenile probation and parole* (pp. 295-315). Monsey, NY: Criminal Justice Press.

Ballucci, D. (2012). Subverting and negotiating risk assessment: A case study of the LSI in a Canadian youth custody facility. *Canadian Journal of Criminology and Criminal Justice, 54,* 203-228.

Bechtel, K., Lowenkamp, C. T., & Latessa, E. (2007). Assessing the risk of re-offending for juvenile offenders using the Youth Level of Service/Case Management Inventory. *Journal of Offender Rehabilitation, 45,* 85-108.

Blase, K., & Fixsen, D. (2013). *Core intervention components: Identifying and operationalizing what makes programs work*. Washington, DC: Department of Health & Human Services. Retrieved January 2, 2014 from the National Implementation Research Network: https://www.nirn.fpg.unc.edu/resources.

Bonett, D. G. (2002). Sample size requirements for estimating intraclass correlations with desired precision. *Statistics in Medicine, 21,* 1331-1335.

Bonta, J. (1996). Risk-needs assessment and treatment. In A. Harland (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply.* (pp.18-32). Thousand Oaks, CA: Sage Publications.

Bonta, J., & Andrews, D. A. (2007). *Risk-Need-Responsivity model for offender assessment and rehabilitation 2007-06*. Retrieved September 28, 2008 from https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/rsk-nd-rspnsvty/index-eng.aspx.

Borum, R., Bartel, P. A., & Forth, A. (2006). *The Structured Assessment of Violence Risk in Youth.* Lutz, FL: Psychological Assessment Resources Incorporated.

Borum, R., Lodewijks, H., Bartel, P. A., & Forth, A. E. (2009). Structured Assessment of Violence Risk in Youth. In R. K. Otto & K. S. Douglas (Eds.), *Handbook of violence risk assessment.* (pp. 63-79). New York, NY: Taylor & Francis Group.

Bouffard, J. A., & Bergseth, K. J. (2008). The impact of re-entry services on juvenile offenders' recidivism. *Youth Violence and Juvenile Justice, 6,* 295-318.

Brame, R., Bushway, S. D., & Paternoster, R. (2004). Assessing the effect of adolescent employment on involvement in criminal activity. *Journal of Contemporary Criminal Justice, 20,* 236-256.

Brock, L., Burrell, J., & Tulipano, T. (2006). *NDTAC issue brief: Family involvement*. Washington, DC: U.S. Department of Education. Retrieved June 16, 2008 from https://www.neglected-delinquent.org/nd/resources/spotlight/spotlight200608b.asp.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Bruns, E. J. & Walker, J. S. (Eds.), (2008). *The resource guide to wraparound*. Portland, OR: National Wraparound Initiative, Research and Training Center for Family Support and Children's Mental Health. Retrieved December 27, 2013 from https://www.nwi.pdx.edu/NWI-book/index.shtml.

Bullis, M., & Yovanoff, P. (2006). Idle hands: Community employment experiences of formerly incarcerated youth. *Journal of Emotional and Behavioral Disorders, 14*, 71-85.

Bullis, M., Yovanoff, P., Mueller, G., & Havel, E. (2002). Life on the 'outs'--examination of the facility-to-community transition of incarcerated youth. *Exceptional Children, 69*, 7-22.

Burl, J. D. (2013). *Are file review-based SAVRY ratings of violence risk reliable?* Unpublished doctoral dissertation, Drexel University, Philadelphia, PA.

Catchpole, R. E. H., & Gretton, H. M. (2003). The predictive validity of risk assessment with violent young offenders: A 1-year examination of criminal outcome. *Criminal Justice and Behavior, 30*, 688-708.

Chung, H. L., Little, M., & Steinberg, L. (2005). The transition to adulthood for adolescents in the juvenile justice system: A developmental perspective. In D. W. Osgood, E. M. Foster, C. Flanagan, & G. R. Ruth (Eds.), *On your own without a net: The transition to adulthood for vulnerable populations.* (pp. 68-91). Chicago, IL US: University of Chicago Press.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6,* 284-290.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Costello, M. Y., & MacCrae-Krisa, L. (2011). *An environmental scan of best practices in supporting persistent youth offenders in the transition to adulthood*. Calgary, AB: Canadian Research Institute for Law and the Family. Retrieved January 2, 2012 from https://www.worldcat.org.

Cottle, C. C., Lee, R. J., & Heilbrun, K. (2001). The prediction of criminal recidivism in juveniles: A meta-analysis. *Criminal Justice and Behavior, 28*, 367-394.

Cunningham, A. (2002). *Lessons learned from a randomized study of Multisystemic Therapy in Canada*. London, ON: Research from the Centre for Children & Families in the Justice System. Retrieved November 14, 2008 from https://www.lfcc.on.ca/onestep.html.

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review, 18,* 23-45.

Department of Corrective Services. (2007). *Intensive Supervision Program.* Perth, WA: Government of Western Australia. Retrieved September 24, 2009 from https://www.correctiveservices.wa.gov.au.

Department of Justice Canada. (2013). *Recent changes to Canada's youth justice system*. Ottawa, ON: Department of Justice Canada. Retrieved December 20, 2013 from https://www.justice.gc.ca/eng/cj-jp/yj-jj/ycja-lsjpa/sheets-feuillets/amend-modif.html.

Dobson, D., & Cook, T. J. (1980). Avoiding Type III error in program evaluation: Results from a field experiment. *Evaluation and Program Planning, 3,* 269-276.

Dowden, C., & Andrews, D. (1999). What works in young offender treatment: A meta-analysis. *FORUM on Corrections Research*, *11*, 21-24.

Dowden, C., & Andrews, D. A. (2003). Does family intervention work for delinquents? Results of a meta-analysis. *Canadian Journal of Criminology and Criminal Justice, 45,* 327-342.

Drake, E. K., & Barnoski, R. (2006). *Recidivism findings for the Juvenile Rehabilitation Administration's mentoring program: Final report.* Olympia, WA: Washington State Institute for Public Policy. Retrieved June 12, 2014 from https://www.wsipp.wa.gov.

Eddy, J. M.,& Chamberlain, P. (2000). Family management and deviant peer association as mediators of the impact of treatment on youth antisocial behavior. *Journal of Consulting and Clinical Psychology*, *68*, 857–863.

Ellis, T., Pamment, N., & Lewis, C. (2009). Public protection in youth justice? The Intensive Supervision and Surveillance Programme from the inside. *International Journal of Police Science & Management, 11,* 393-413.

Ferguson, J. L., (2002). Putting the "What works" research into practice: An organizational perspective. *Criminal Justice and Behavior, 29,* 472-492.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M. & Wallace, F. (2005). *Implementation research: A synthesis of the literature* (FMHI Publication #231). Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute. Retrieved January 2, 2014 from the National Implementation Research Network: https://www.nirn.fpg.unc.edu/resources.

Flores, A. W., Russell, A. L., Latessa, E. J., & Travis, L. F. (2005). Evidence of professionalism or quackery: Measuring practitioner awareness of risk/need factors and effective treatment strategies. *Federal Probation, 69,* 9-14.

Flores, A. W., Travis, L. F., & Latessa, E. J. (2004). *Case classification for juvenile corrections: An assessment of the Youth Level of Service/ Case Management Inventory (YLS/CMI), Final report*. Retrieved February 6, 2014 from https://www.ncjrs.gov/pdffiles1/nij/grants/204005.pdf.

Forgatch, M. S., Patterson, G. R., & DeGarmo, D. S. (2005). Evaluating fidelity: Predictive validity for a measure of competent adherence to the Oregon model of Parent Management Training. *Behavior Therapy, 36,* 3-13.

Gammelgård, M., Koivisto, A., Eronen, M., & Kaltiala-Heino, R. (2008). The predictive validity of the structured assessment of violence risk in youth (SAVRY) among institutionalised adolescents. *Journal of Forensic Psychiatry & Psychology, 19*, 352-370.

Gendreau, P. (1996). The principles of effective intervention with offenders. In F. X. Harland, (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply* (pp. 117-130). Thousand Oaks, CA: Sage.

Gendreau, P., Goggin, C., Cullen, F.T., & Andrews, D.A. (2000). The effects of community sanctions and incarceration on recidivism. *Forum on Corrections Research, 12,* no pagination. Retrieved July 30, 2011 from https://www.csc-scc.gc.ca/text/pblct/forum/e122/e122c-eng.shtml.

Gendreau, P., Goggin, C., & Paparozzi, M. (1996). Principles of effective assessment for community corrections. *Federal Probation, 60*, 64-70.

Gendreau, P., Goggin, C., & Smith, P. (1999). The forgotten issue in effective correctional treatment: Program implementation. *International Journal of Offender Therapy and Comparative Criminology, 43,* 180-187.

Gies, S.V. (2003). Aftercare services. *Juvenile Justice Practice Series*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention. Retrieved June 7, 2008 from https://www.ncjrs.gov/html/ojjdp/201800/contents.html.

Goodstein, L., & Sontheimer, H. (1997). The implementation of an intensive aftercare program for serious juvenile offenders: A case study. *Criminal Justice and Behavior*, *24*, 332-359.

Gray, E., Taylor, E., Roberts, C., Merrington, S., Fernandez, R., & Moore, R. (2005). *Intensive Supervision and Surveillance Programme: The final report.* University of Oxford, London: Youth Justice Board for England and Wales. Retrieved August 19, 2008 from https://yjbpublications.justice.gov.uk/en-gb/Resources/Downloads/ ISSP%20the%20final%20report.pdf.

Greenwood, P. W., Deschenes, E. P., & Adams, J. (1993). *Chronic juvenile offenders: Final results from the Skillman Aftercare Experiment*. Santa Monica, CA: Rand. Retrieved June 22, 2008 from https://www.rand.org.

Gresham, F. M., Gansle, K. A., & Noell, G. H. (1993). Treatment integrity in Applied Behavior Analysis with children. *Journal of Applied Behavior Analysis, 26,* 257-263.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and analysis. *Tutorials in Quantitative Methods for Psychology, 8,* 23-34.

Hannah-Moffat, K., & Maurutto, P. (2003). *Youth risk/need assessment: An overview of issues and practice.* Ottawa, ON: Youth Justice Policy, Research and Statistics Division, Department of Justice Canada. Retrieved January 11, 2008 from https://www.justice.gc.ca/ eng/pi/yj-jj/res-rech/resrech.html.

Harris, P.W., Lockwood, B., & Mengers, L. (2009). *A CJCA white paper: Defining and measuring recidivism* [White paper]. Braintree, MA: Council of Juvenile Correctional Administrators. Retrieved January 9, 2014 from https://www.cjca.net.

Hawkins, S. R., Dawes, D., Lattimore, P. K., & Visher, C. A. (2009). *Re-entry experiences of confined juvenile offenders: Characteristics, service receipt, and outcomes of juvenile male participants in the SVORI multi-site evaluation.* Retrieved January 14, 2014 from https://www.svori-evaluation.org/.

Henggeler, S. W. (1999). Multisystemic Therapy: An overview of clinical procedures, outcomes, and policy implications. *Child Psychology and Psychiatry Review, 4,* 2-10.

Henggeler, S. W., & Sheidow, A. J. (2012). Empirically supported family-based treatments for conduct disorder and delinquency in adolescents. *Journal of Marital and Family Therapy, 38,* 30-58.

Herrenkohl, T. I., Maguin, E., Hill, K. G., Hawkins, J. D., Abbott, R. D., & Catalano, R. F. (2000). Developmental risk factors for youth violence. *Journal of Adolescent Health, 26,* 176-186.

Hill, J. K., Andrews, D. A., & Hoge, R. D. (1991). Meta-analysis of treatment programs for young offenders: The effect of clinically relevant treatment on recidivism, with controls for various methodological variables. *The Canadian Journal of Program Evaluation, 6,* 97-109.

Hoge, R. D. (2002). Standardized instruments for assessing risk and need in youthful offenders. *Criminal Justice and Behavior, 29*, 380-396.

Hoge, R., D. (2005). Youth Level of Service/Case Management Inventory. In T. Grisso, G. Vincent, & D. Seagrave (Eds.), *Mental health screening and assessment in juvenile justice.* (pp. 283-294). New York, NY: Guilford Press.

Hoge, R. D., & Andrews, D. A. (2011). *Youth Level of Service/Case Management Inventory 2.0.* Toronto, ON: Multi-health Systems Incorporated.

Hoge, R. D., Andrews, D. A., & Leschied, A. W. (1996). An investigation of risk and protective factors in a sample of youthful offenders. *Journal of Child Psychology and Psychiatry, 37*, 419-424.

Iutcovich, J. M., & Pratt, D. J. (1998). *A final report for: Assessment of aftercare services provided to delinquent youth*. Erie, PA: Keystone University Research Corporation. Retrieved June 3, 2008 from https://www.portal.state.pa.us/portal/ server.pt/document/882096/aftercare_combinedcontent_pdf.

Jack, L. (2000). *Psychopathy*, *risk/need factors*, *and psychiatric symptoms in high-risk youth: Relationships between variables and their link to recidivism.* Unpublished doctoral dissertation, Simon Fraser University, Burnaby, BC.

James, C., Stams, G. J. J. M., Asscher, J. J., De Roo, A. K., & van der Laan, P. H. (2013). Aftercare programs for reducing recidivism among juvenile and young adult offenders: A meta-analytic review. *Clinical Psychology Review, 33,* 263-274.

Jung, S., & Rawana, E. P. (1999). Risk and need assessment of juvenile offenders. *Criminal Justice and Behavior, 26,* 69-89.

Juvenile Justice New South Wales. (2012). *Intensive Supervision Program*. Retrieved December 20, 2013 from https://www.lawlink.nsw.gov.au.

Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research, 30,* 411-433.

Krohn, M. D., Thornberry, T. P., Gibson, C. L., & Baldwin, J. M. (2010). The development and impact of self-report measures of crime and delinquency. *Journal of Quantitative Criminology, 26,* 509-525.

Krysik, J., & LeCroy, C. W. (2002). The empirical validation of an instrument to predict risk of recidivism among juvenile offenders. *Research on Social Work Practice, 12*, 71-81.

Kumpfer, K. L. (1993). *Strengthening America's families: Promising parenting strategies for delinquency prevention-user's guide Office of Juvenile Justice and Delinquency Prevention*. Office of Juvenile Programs. U.S. Department of Justice, Washington, D.C.

Lane, J., Turner, S., Fain, T., & Sehgal, A. (2005). Evaluating an experimental intensive juvenile probation program: Supervision and official outcomes. *Crime & Delinquency, 51,* 26-52.

Lane, J., Turner, S., Fain, T., & Sehgal, A. (2007). The effects of an experimental intensive juvenile probation program on self-reported delinquency and drug use. *Journal of Experimental Criminology, 3,* 201-219.

Land, K. C., McCall, P. L., & Williams, J. R. (1990). Something that works in juvenile justice: An evaluation of the North Carolina Court Counselors' Intensive Protective Randomized Experimental Project, 1987-1989. *Evaluation Review*, *14*, 574-606.

Långström, N., Grann, M., Tengström, A., Lindholm, N., Woodhouse, A., & Kullgren, G. (1999). Extracting data in file-based forensic psychiatric research: Some methodological considerations. *Nordic Journal of Psychiatry, 53*, 61-67.

Latimer, J. (2001). A meta-analytic investigation of youth delinquency, family treatment, and recidivism. *Canadian Journal of Criminology, 43,* 237-253.

Latimer, J., Dowden, C., & Morton-Bourgon, K. E. (2003). *Treating youth in* conflict *with the law: A new meta-analysis.* Ottawa, ON: Research and Statistics Division, Department of Justice Canada. Retrieved February 18, 2009 from https://www.justice.gc.ca/eng/rp-pr/cj-jp/yj-jj/rr03_yj3-rr03_jj3/rr03_yj3.pdf.

Lattimore, P. K., Brumbaugh, S., Visher, C., Lindquist, C., Winterfield, L., Salas, M., & Zweig, J. (2004). *National portrait of the Serious and Violent Offender Reentry Initiative.* Washington, DC: Urban Institute. Retrieved June 16, 2008 from https://www.svori-evaluation.org/.

Lattimore, P. K., Visher, C. A., & Linster, R. L. (1995). Predicting rearrest for violence among serious youthful offenders. *Journal of Research in Crime and Delinquency, 32*, 54-83.

Lattimore, P. K., & Visher, C. A. (2010). *The multi-site evaluation of SVORI: Summary and synthesis.* Retrieved January 14, 2014 from https://www.svori-evaluation.org/.

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about inter-rater reliability and inter-rater agreement. *Organizational Research Methods, 11,* 815-852.

Lidz, C. W., Mulvey, E. P., & Gardner, W. (1993). The accuracy of predictions of violence to others. *Journal of the American Medical Association, 269,* 1007-1111.

Lipsey, M. W. (1992). The effect of treatment on juvenile delinquents: Results from meta-analysis. In F. Lösel, D. Bender, & T. Bliesener (Eds.), *Psychology and law: International perspectives.* (pp. 131-143). Oxford England: Walter De Gruyter.

Lipsey, M. W. (1995). What do we learn from 400 research studies on the effectiveness of treatment with juvenile delinquents? In J. McGuire (Ed.), *What works: Reducing reoffending: Guidelines from research and practice.* (pp. 63-78). Oxford England: John Wiley & Sons.

Lipsey, M. W. (1999). Can rehabilitative programs reduce the recidivism of juvenile offenders? An inquiry into the effectiveness of practical programs. *Virginia Journal of Social Policy & the Law*, *6*, 611–641.

Lipsey, M. W., & Derzon, J. H. (1998). Predictors of violent or serious delinquency in adolescence and early adulthood: A synthesis of longitudinal research. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 86–105). Thousand Oaks, CA: Sage.

Lipsey, M. W., Howell, J. C., Kelly, M. R., Chapman, G., & Carver, D. (2010). *Improving the effectiveness of juvenile justice programs: A new perspective on evidence based practice*. Georgetown University: Center for Juvenile Justice Reform. Retrieved January 13, 2014 from https://cjjr.georgetown.edu// resources2/cjjrpublications/cjjr_publications.html

Lipsey, M. W., & Wilson, D. B. (1998). Effective intervention for serious juvenile offenders: A synthesis of research. In R. Loeber & D. P. Farrington (Eds.), *Serious & violent juvenile offenders: Risk factors and successful interventions.* (pp. 313-345). Thousand Oaks, CA US: Sage Publications, Inc.

Listwan, S. J., Cullen, F. T., & Latessa, E. J. (2006). How to prevent prisoner re-entry programs from failing: Insights from evidence-based corrections. *Federal Probation, 70,* 19-25.

Lodewijks, H. P. B., de Ruiter, C., & Doreleijers, T. A. H. (2008). Gender differences in violent outcome and risk assessment in adolescent offenders after residential treatment. *The International Journal of Forensic Mental Health, 7*, 133-146.

Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage Publications.

Loughran, T. A., Mulvey, E. P., Schubert, C. A., Fagan, J., Piquero, A. R., & Losoya, S. H. (2009). Estimating a dose-response relationship between length of stay and future recidivism in serious juvenile offenders. *Criminology: An Interdisciplinary Journal, 47*, 699-740.

Lowenkamp, C. T., & Latessa, E. J. (2004). Understanding the risk principle: How and why correctional interventions can harm low-risk offenders. In *Topics in community corrections: Assessment issues for Managers* (pp. 3–8). Washington, DC: Department of Justice. Retrieved May 10, 2014 from https://www.nicic.gov/library/.

Lowenkamp, C. T., Latessa, E. J., & Smith, P. (2006). Does correctional program quality really matter? The impact of adhering to the principles of effective intervention. *Criminology & Public Policy, 5,* 575-594.

Lowenkamp, C. T., Makarios, M. D., Latessa, E. J., Lemke, R., & Smith, P. (2010). Community corrections facilities for juvenile offenders in Ohio: An examination of treatment integrity and recidivism. *Criminal Justice and Behavior, 37,* 695-708.

Luong, D., & Wormith, J. S. (2011). Applying risk/need assessment to probation practice and its impact on the recidivism of young offenders. *Criminal Justice and Behavior, 38,* 1177-1199.

Mackenzie, D. L. (2006). *What works in corrections: Reducing the criminal activities of offenders and delinquents.* New York, NY: Cambridge University Press.

Marczyk, G. R., Heilbrun, K., Lander, T., & DeMatteo, D. (2003). Predicting juvenile recidivism with the PCL:YV, MAYSI, and YLS/CMI. *The International Journal of Forensic Mental Health, 2*, 7-18.

McCreary Centre Society. (2005). *Time out II: A profile of BC Youth in custody.* Vancouver, BC: McCreary Centre Society. Retrieved November 26, 2007 from McCreary Centre Society: https://www.mcs.bc.ca.

McCreary Centre Society. (2012). *PLEA evaluation report: PLEA programs for youth in conflict with the law.* Vancouver, BC: McCreary Centre Society. Retrieved December 23, 2013 from https://www.mcs.bc.ca.

McCreary Centre Society. (2014). *Voices from the Inside II: A Next Steps project with youth in custody*. Vancouver, BC: McCreary Centre Society. Retrieved January 21, 2014 from https://www.mcs.bc.ca.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1,* 30-46.

Mihalic, S., Fagan, A., Irwin, K., Ballard, D., & Elliott, D. (2004). *Blueprints for violence prevention.* Washington, DC: Office of Juvenile Justice and Delinquency Prevention. Retrieved December 8, 2008 from https://www.ncjrs.gov/pdffiles1/ojjdp/187079.pdf.

Mihalic, S. F., & Irwin, K. (2003). Blueprints for Violence Prevention: From research to real-world settings – factors influencing the successful replication of model programs. *Youth Violence and Juvenile Justice, 1,* 307-329.

Mihalic, S., Irwin, K., Fagan, A., Ballard, D., & Elliott, D. (2004). Successful program implementation: Lessons learned from Blueprints. *Juvenile Justice Bulletin, July,* 1-13.

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding Analysis of Covariance. *Journal of Abnormal Psychology, 110,* 40-48.

Ministry of Children and Family Development. (2006). *Guidelines for Intensive Support and Supervision Programs*. (unpublished document).

Ministry of Children and Family Development. (2012). BC enhancing youth custody services [Press release]. Retrieved December 20, 2013 from https://www2.news.gov.bc.ca/news_releases_2009-2013/2012CFD0002-000038.htm.

Ministry of Children and Youth Services Youth Justice Services Division. (2011). Ontario's community-based mental health responses for youth in conflict with the law. Paper presented at the Canadian Criminal Justice Association Pan-Canadian Congress, QC.

Moore, R., Gray, E., Roberts, C., & Taylor, E. (2004). *ISSP: The Initial Report on the Intensive Supervision and Surveillance Programme.* University of Oxford, London: Youth Justice Board for England and Wales. Retrieved August 19, 2008 from https://yjbpublications.justice.gov.uk/publications/Resources/Downloads/ISSP_summary.pdf.

Mordell, S., van der Woerd, K., Viljoen, J., Roesch, R., & Penner, E. (2008). *Intensive Support and Supervision Program best practices review.* Internal report prepared for British Columbia Youth Custody Services, Ministry of Children and Family Development.

Mossman, D. (1994). Assessing predictions of accuracy: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology, 62,* 783-792.

Nation, M., Crusto, C., Wandersman, A., Kumpfer, K. L., Seybolt, D., Morrissey-Kane, E., & Davino, K. (2003). What works in prevention: Principles of effective prevention programs. *American Psychologist, 58,* 449-456.

Office of Juvenile Justice and Delinquency. (2004). A message from OJJDP. *Juvenile Justice Bulletin, July,* 1.

Office of Juvenile Justice and Delinquency. (n.d.). *Model programs guide*. Retrieved January 15, 2014 from http://www.ojjdp.gov/mpg/Program.

Olver, M. E., Stockdale, K. C., & Wong, S. C. P. (2012). Short and long-term prediction of recidivism using the Youth Level of Service-Case Management Inventory in a sample of serious young offenders. *Law and Human Behavior, 36,* 331-344.

Onifade, E., Davidson, W., Campbell, C., Turke, G., Malinowski, J., & Turner, K. (2008). Predicting recidivism in probationers with the youth level of service case management inventory (YLS/CMI). *Criminal Justice and Behavior, 35*, 474-483.

Paparozzi, M. A., & Gendreau, P. (2005). An intensive supervision program that worked: Service delivery, professional orientation, and organizational supportiveness. *The Prison Journal, 85*, 445-466.

Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology, 49*, 1373-1379.

Perrault, R. T., Paiva-Salisbury, M., & Vincent, G. M. (2012). Probation officers' perceptions of youths' risk of reoffending and use of risk assessment in case management. *Behavioural Sciences and the Law, 30,* 487-505.

Petersilia, J., & Turner, S. (1991). An evaluation of intensive probation in California. *Northwestern School of Law Journal of Criminal Law & Criminology, 82*, 610 – 658.

Petrosino, A., Turpin-Petrosino, C., Hollis-Peel, M., & Lavenberg, J.G. (2013). Scared Straight and other juvenile awareness programs for preventing juvenile delinquency: A systematic review. *Campbell Systematic Reviews, 9.* Retrieved December 20, 2013 from https://www.campbellcollaboration.org/lib/project/3/.

Random.org. (n.d.). True random number generator. In *Random.org.* Retrieved May 10, 2012 from https://www.random.org/.

Representative for Children and Youth, & Office of the Provincial Health Officer. (2009). *Health and wellbeing of children in care: Youth justice experiences and outcomes.* Retrieved August 12, 2011 from https://www.rcybc.ca/Images/PDFs/Reports/Youth%20Justice%20Joint%20Rpt%20 FINAL%20.pdf.

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC area, Cohen's *d,* and *r*. *Law and Human Behavior, 29,* 615–620.

Roth, P. L. (1994). Missing data: A conceptual review for applied psychologists. *Personnel Psychology, 47,* 537-560.

Rubin, D. B. (1987). *Multiple imputation for non-response in surveys.* New York, NY: Wiley.

Saunders, S. M., Resnick, M. D., Hoberman, H. M., & Blum, R. W. (1994). Formal help-seeking behaviour of adolescents identifying themselves as having mental health problems. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 718-728.

Schmidt, F., Campbell, M., & Houlding, C. (2011). Comparative analyses of the YLS/CMI, SAVRY, and PCL:YV for adolescent offenders: A 10 year follow-up into adulthood. *Youth Violence and Juvenile Justice, 9,* 23-42.

Schmidt, F., Hoge, R. D., & Gomes, L. (2005). Reliability and validity analyses of the Youth Level of Service/Case Management Inventory. *Criminal Justice and Behavior, 32,* 329-344.

Sherman, L. W., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P., & Bushway, S. (1997). *Preventing crime: What works, what doesn't, what's promising*. Washington, D.C.: United States Department of Justice. Retrieved September 3, 2011 from https://www.ncjrs.gov/pdffiles/171676.PDF.

Shook, J. J., & Sarri, R. C. (2007). Structured decision making in juvenile justice: Judges' and probation officers' perceptions of use. *Children and Youth Services, 29,* 1335-1351.

Shoukri, M. M. (2011). *Measures of interobserver agreement and reliability,* 2nd ed. New York, NY: CRC Press.

Shoukri, M. M., Asyali, M. H., & Donner, A. (2004). Sample size requirements for the design of a reliability study: Review and new results. *Statistical Methods in Medical Research, 13,* 251-271.

Shrout, P. E., & Fleiss, J. L., (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420-428.

Siegel, S., & Castellan, N.J. (1988). *Nonparametric statistics for the behavioral sciences* (2nd Ed). New York, NY: McGrawHill.

Simons, K. J., Paternite, C. E., & Shore, C. (2001). Quality of parent/adolescent attachment and aggression in young adolescents. *The Journal of Early Adolescence, 21*, 182-203.

Soper, D. (n.d.). A-priori sample size calculator for hierarchical multiple regression. In *Statistics Calculators, version 3.0.* Retrieved June 28, 2012 from https://danielsoper.com/statcalc3/calc.aspx?id=16.

Staff, J., Osgood, D. W., Schulenberg, J. E., Bachman, J. G., & Messersmith, E. E. (2010). Explaining the relationship between juvenile delinquency and employment. *Criminology, 48,* 1101-1131.

Stouthamer-Loeber, M., Wei, E., Loeber, R., & Masten, A. S. (2004). Desistance from persistent serious delinquency in the transition to adulthood. *Development and Psychopathology, 16,* 897-918.

Spencer, M. B., & Jones-Walker, C. (2004). Interventions and services offered to former juvenile offenders reentering their communities: An analysis of program effectiveness. *Youth Violence and Juvenile Justice, 2*, 88-97.

Statistics Canada. (n.d.). *Table 252-0050 Court, youth cases by type of sentence: intensive support and supervision, probation, attendance at non-residential program, fine, by province and territory 2009/2010* (table). CANSIM (database). Last updated May 4, 2011. Retrieved July 8, 2011 from https://www40.statcan.ca/l01/cst01/legal43a-eng.htm.

Statistics Canada. (n.d.). *Cases completed in youth court, by type of offence and selected sentence, Canada, 2010/2011.* Last updated May 31, 2013. Retrieved January 21, 2014 from https://www.statcan.gc.ca/pub/85-002-x/2012001/article/11645/tbl/tbl05-eng.htm.

Steinberg, L. (2009). Adolescent development and juvenile justice. *Annual Review of Clinical Psychology, 5,* 459-485.

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G…Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. *British Medical Journal, 338,* 157-160.

Stillman, A. R., Hadley-Ives, E., Doré, P., Polgar, M., Horvath, V. E., Striley, C., & Elze, D. (2000). Youth's access to mental health services: The role of providers' training resource connectivity, and assessment of need. *Mental Health Services Research, 2,* 141-154.

Sukhodolsky, D. G., & Ruchkin, V. (2006). Evidence-based psychosocial treatments in the juvenile justice system. *Child and Adolescent Psychiatric Clinics of North America, 15*, 501.

Suter, J. C. & Bruns, E. J. (2009). Effectiveness of the wraparound process for children with emotional and behavioral disorders: A meta-analysis. *Clinical Child and Family Psychology Review, 12,* 336-351.

Thompson, A. P., & Pope, Z. (2005). Assessing juvenile offenders: Preliminary data for the Australian adaptation of the youth level of Service/Case management inventory (Hoge & Andrews, 1995). *Australian Psychologist, 40*, 207-214.

Tolan, P., Henry, D., Schoeny, M., Bass, A, Lovegrove, P., & Nichols, E. (2013). Mentoring interventions to affect juvenile delinquency and other problems: A systematic review. *Campbell Systematic Review, 10.* Retrieved January 8, 2014 from https://www.campbellcollaboration.org/lib.

Trulson, C. R., Marquart, J. W., Mullings, J. L., & Caeti, T. J. (2005). In between adolescence and adulthood: Recidivism outcomes of a cohort of state delinquents. *Youth Violence and Juvenile Justice, 3*, 355-387.

Unruh, D., Bullis, M., & Yovanoff, P. (2003). Community reintegration outcomes for formerly incarcerated adolescent fathers and non-fathers. *Journal of Emotional and Behavioral Disorders, 11*, 144-156.

Unruh, D. K., Gau, J. M., & Waintrup, M. G. (2009). An exploration of factors reducing recidivism rates of formerly incarcerated youth with disabilities participating in a re-entry intervention. *Journal of Child and Family Studies, 18,* 284-293.

van der Put, C. E., Stams, G. J. J., Hoeve, M., Dekovic, M., Spanjaard, H. J. M., van der Laan, P. H., & Barnoski, R. P. (2012). Changes in the relative importance of dynamic risk factors for recidivism during adolescence. *International Journal of Offender Therapy and Comparative Criminology, 56,* 296-316.

van der Put, C. E., van Vugt, E. S., Stams, G. J. J. M., Dekovic, M., & van der Laan, P. H. (2013). Differences in the prevalence and impact of risk factors for general recidivism between different types of juveniles who have committed sexual offences (JSOs) and juveniles who have committed nonsexual offences (NSOs). *Sexual Abuse: A Journal of Research and Treatment, 25,* 41-68.

van Marle, Hjalmar J. C., Hempel, I. S., & Buck, N. M. L. (2010). Young serious and vulnerable offenders in the Netherlands: A cohort follow-up study after completion of a PIJ (detention) order. *Criminal Behaviour and Mental Health, 20*, 349-360.

Van Voorhis, P., & Brown, K. (1997). *Risk classification in the 1990s* (draft). Retrieved February 6, 2014 from https://nicic.gov/Library/013243.

Vancouver Coastal Youth Justice Services. (2008). *Annual program report 2007/2008*. Vancouver, BC: PLEA Community Services Society of BC. Retrieved November 14, 2008 from https://www.plea.ca/studies-reports.

Vieira, T. A., Skilling, T. A., & Peterson-Badali, M. (2009). Matching court-ordered services with treatment needs: Predicting treatment success with young offenders. *Criminal Justice and Behavior, 36*, 385-401.

Viljoen, J. L., Elkovitch, N., Scalora, M. J., & Ullman, D. (2009). Assessment of reoffence risk in adolescents who have committed sexual offences: Predictive validity of the ERASOR, PCL:YV, YLS/CMI, and static-99. *Criminal Justice and Behavior, 36*, 981-1000.

Viljoen, J. L., Scalora, M., Ullman, D., Cuadra, L., Bader, S., Chavez, V., & Lawrence, L. (2008). Assessing risk for violence in adolescents who have sexually offended: A comparison of the J-SOAP-II, SAVRY, and J-SORRAT-II. *Criminal Justice and Behavior, 35*, 5-23.

Vincent, G. M., Chapman, J., & Cook, N. E. (2011). Risk-needs assessment in juvenile justice: Predictive validity of the SAVRY, racial differences, and the contribution of needs factors. *Criminal Justice and Behavior, 38,* 42-62.

Vincent, G. M., Guy, L. S., Gershenson, B. G., & McCabe, P. (2012). Does risk assessment make a difference? Results of implementing the SAVRY in juvenile probation. *Behavioral Sciences and the Law, 30,* 384-405.

Vincent, G. M., Paiva-Salisbury, M. L., Cook, N. E., Guy, L. S., & Perrault, R. T. (2012). Impact of risk/needs assessment on juvenile probation officers' decision making: Importance of implementation. *Psychology, Public Policy, and Law, 18,* 549-576.

Walker, J. A., & Friedman, K. (2001). *Listening and learning from families in juvenile justice*. Columbia, MD: Maryland Coalition of Families for Children's Mental Health. Retrieved June 16, 2008 from https://www.mdcoalition.org/images/ stories/publications/jjustice1.pdf.

Walter, S. D., Eliasziw, M., & Donner, A. (1998). Sample size and optimal designs for reliability studies. *Statistics in Medicine, 17,* 101-110.

Ward-Horner, J., & Sturmey, P. (2010). Component analyses using single-subject experimental designs: A review. *Journal of Applied Behavioral Analysis, 43,* 685-704.

Welsh, J. L., Schmidt, F., McKinnon, L., Chattha, H. K., & Meyers, J. R. (2008). A comparative study of adolescent risk assessment instruments: Predictive and incremental validity. *Assessment, 15*, 104-115.

Wiebush, R. G., Baird, C., Krisberg, B., & Onek, D. (1995). Risk assessment and classification for serious, violent, and chronic juvenile offenders. In J. C. Howell, B. F. Krisberg, J. D. Hawkins, &J. J. Wilson (Eds.), *A sourcebook: Serious, violent, and chronic juvenile offenders*. (pp. 171-212). Thousand Oaks, CA: Sage.

Wiebush, R. G., Wagner, D., McNulty, B., Wang, Y., & Le, T. N. (2005). *Implementation and outcome evaluation of the intensive aftercare program: Final report.* Washington, DC: Office of Juvenile Justice and Delinquency Prevention. Retrieved June 3, 2008 from https://www.ncjrs.gov/app/publications/abstract.aspx?ID=206177.

Wilson, J. A., & Davis, R. C. (2006). Good intentions meet hard realities: An evaluation of the Project Greenlight re-entry program. *Criminology & Public Policy, 5,* 303-338.

Wilson, S. J., & Lipsey, M.W. (2005). *The effectiveness of school-based violence prevention programs for reducing disruptive and aggressive behavior*. Retrieved January 7, 2014 from https://www.ncjrs.gov/pdffiles1/nij/grants/211376.pdf.

Youth Criminal Justice Act (2006). Retrieved July 8, 2011 from the Department of Justice Canada Website: https://laws-lois.justice.gc.ca/eng/acts/Y-1.5/page-1.html.

Ziskin, J. (1995). Challenging clinical judgment. In J. Ziskin (Ed.), *Coping with psychiatric and psychological evidence* (5th ed., pp. 195–292). Los Angeles, CA: Law & Psychology Press.

Appendix A.

Tables and Figures

Table 1. Demographic Characteristics of Excluded Youth and Group Differences from Study Sample

		Excluded youth			
	Total (*n* = 314)	**MCAR** (*n* = 14)	**MAR** (*n* = 97)	**MNAR** (*n* = 27)	**Sample** (*n* = 176)
Age					
M(SD)	16.31(1.23)	16.80(0.82)	16.13(1.32)	16.59(1.91)	16.32(1.20)
		t(188)=1.47	*t*(271)=-1.20	*t*(201)=1.09	
Gender					
% Male	80.9	78.6	85.6	92.6	76.7
% Female	19.1	21.4	14.4	7.4	23.3
		$\chi^2(1)=0.02$	$\chi^2(1)=3.05$	$\chi^2(1)=3.54$	
Aboriginal Identification					
% Any Aboriginal	25.8	21.4	34.0	18.5	22.7
% No Aboriginal	33.8	50.0	35.1	25.9	33.0
% Unknown	40.4	28.6	30.9	55.6	44.3
		$\chi^2(2)=1.86$	$\chi^2(2)=5.90$	$\chi^2(2)=1.19$	
Location					
% Burnaby	67.2	78.6	71.1	85.2	61.4
% Prince George	9.2	7.1	3.1	0.0	13.6
% Victoria	23.6	14.3	25.8	14.8	25.0
		$\chi^2(2)=2.36$	$\chi^2(2)=7.96$*	$\chi^2(2)=6.77$*	

Note. Excluded youth were grouped according to: MCAR = Missing completely at random, MAR = Missing at random, MNAR = Missing not at random.

Missing groups were contrasted with the study sample using an independent samples *t*-test for age and Pearson's chi-square for categorical variables.

* significantly different from study sample at $p < 0.05$.

Table 2. Offence Characteristics of Excluded Youth and Group Differences from Sample

		Excluded Youth			
	Total (n = 314)	**MCAR** (n = 14)	**MAR** (n = 97)	**MNAR** (n = 27)	**Sample** (n = 176)
Prior to ISSP					
# of offences	5.97(5.35)	5.93(4.98)	5.81(5.35)	5.69(4.49)	6.11(5.53)
		t(187)=-0.12	*t*(268)=-0.43	*t*(199)=-0.37	
# of violent	1.94(2.19)	1.64(2.24)	1.76(1.60)	1.85(2.13)	2.07(2.47)
		t(187)=-0.63	*t*(268)=-1.13	*t*(199)=-0.45	
# of non-breach	3.67(3.77)	3.00(2.42)	3.51(3.53)	3.04(2.54)	3.91(4.11)
		t(187)=-0.82	*t*(268)=-0.81	*t*(199)=-1.05	
# of offence dates	4.98(4.11)	5.00(3.84)	4.88(4.10)	4.88(3.90)	5.04(4.19)
		t(187)=-0.04	*t*(268)=-0.29	*t*(199)=-0.18	
ISSP & follow up					
# of offences	3.99(5.36)	3.43(5.10)	4.91(5.94)	1.19(2.00)	3.95(5.27)
		t(187)=-0.36	*t*(269)=1.37	*t*(199)=-4.93**	
# of violent	0.82(1.60)	1.21(2.89)	.95(1.60)	0.23(0.59)	0.80(1.55)
		t(187)=0.89	*t*(269)=0.74	*t*(199)=-3.46**	
# of non-breach	1.92(3.15)	1.71(3.45)	2.22(3.33)	0.54(0.99)	1.98(3.20)
		t(187)=-0.29	*t*(269)=0.59	*t*(199)=-4.64**	
# of offence dates	3.40(4.57)	3.14(4.60)	4.20(5.21)	1.04(1.76)	3.34(4.37)
		t(187)=-0.16	*t*(269)=1.45	*t*(199)=-4.82**	

Note. Excluded youth were grouped according to: MCAR = Missing completely at random, MAR = Missing at random, MNAR = Missing not at random.

Missing group means were compared to the study group with independent samples *t*-tests.

**significantly different from study sample at $p < 0.001$.

Table 3. Final Sample Characteristics

	% of Youth	M(SD)	Median	Min.	Max.
Actual start age	--	16.23(1.23)	16.36	12.63	17.98
Psychosocial					
MH issues – Yes	17.1	--	--	--	--
Addictions – Yes	48.3	--	--	--	--
FASD - ?/Yes	8.1/7.5	--	--	--	--
Low cognitive – Yes	21.1	--	--	--	--
Offences Prior to ISSP					
Convictions	100	6.50(5.93)	4	1	27
Violent	81.2	2.02(2.05)	2	0	16
Indictable	59.1	1.32 (1.72)	1	0	9
Non-breach	98.3	3.86(3.63)	3	0	20
Seriousness score	--	21.56(18.54)	16	1	98
Days in custody	65.3	79.06(145.89)	14	0	764
Only sexual offences	7.4	--	--	--	--
Offences on ISSP & Follow Up					
Convictions	69.3	3.81(4.49)	2	0	17
Violent	34.1	0.64(1.15)	0	0	6
Indictable	21.6	0.63(1.60)	0	0	10
Non-breach	47.7	1.61(2.51)	0	0	12
Seriousness score	--	9.60 (12.90)	4	0	55
Days in custody	55.7	85.68(161.14)	11	0	1132
Offences on Follow Up Only					
Convictions	42.0	1.41(2.33)	0	0	12
Violent	21.0	0.34(0.82)	0	0	5
Indictable	12.5	0.32(1.13)	0	0	9
Non-breach	38.1	0.70(1.42)	0	0	9
Seriousness score	--	4.37(7.89)	0	0	45
Days in custody	30.1	38.91(82.49)	0	0	365
Offences Post-Follow Up					
Convictions	36.9	1.53(3.66)	0	0	24
Violent	11.9	0.22(0.78)	0	0	5
Days in custody	28.5	74.91(179.01)	0	0	1020
Currently justice-involved	36.4	--	--	--	--

Note. Figures differ from sample statistics in Tables 1 and 2 due to inconsistencies between the database and file sources of information.

Table 4. Inter-rater Reliability Data for File Outcome Variables

	n	$ICC_{(1,1)}$	95% CI	Sig.
Continuous Variables				
Fail to report	10	0.68	0.17 – 0.91	$p < 0.01$
Breach violations	10	0.62	0.06 – 0.94	$p < 0.05$
Police contact	9	0.79	0.35 – 0.95	$p < 0.01$
Arrests	10	0.26	-0.37 – 0.74	*ns*
Offending	10	0.25	-0.39 – 0.74	*ns*
Incidents of violence	10	0.92	0.72 – 0.98	$p < 0.001$
Curfew violations	10	0.74	0.28 – 0.93	$p < 0.01$
Unknown whereabouts	10	0.77	0.34 – 0.94	$p < 0.01$
Drug/alcohol use	10	0.69	0.19 – 0.91	$p < 0.01$
# of programs	10	0.71	0.21 – 0.92	$p < 0.01$
Missed program sessions	7	0.80	0.27 – 0.96	$p < 0.01$
Employment overall	10	0.85	0.53 – 0.96	$p < 0.001$
School enrollment overall	10	0.69	0.18 – 0.91	$p < 0.01$
ISSP engagement	7	0.62	-0.10 – 0.92	$p < 0.05$
	n	**K**	**95% CI**	**Sig.**
Dichotomous Variables				
Any part-time employment	10	0.31	-0.12 – 0.74	*ns*
Any full-time employment	10	0.74	0.20 – 0.99	$p < 0.05$
Any school enrollment	10	0.74	0.20 – 0.99	$p < 0.05$
School at ISSP end	9	0.73	0.17 – 0.99	$p < 0.05$
Employment at ISSP end	9	0.73	0.17 – 0.99	$p < 0.05$

Note. ICC = Intra-class Correlation Coefficient; CI = Confidence Interval; K = Siegel and Castellan's Kappa.

Table 5. Intra-class Correlation Coefficients for Best Practice Variables

	n	$ICC_{(2,1)}$	95% CI	Sig.
ISSP log quality	20	0.78	0.53 – 0.91	$p < 0.001$
File ISSP quality	20	0.62	0.25 – 0.83	$p < 0.01$
Select high-risk youth (BP)	20	0.95	0.89 – 0.98	$p < 0.001$
Select youth by criteria (ISSP)	19	0.87	0.69 – 0.95	$p < 0.001$
Risk assessment for youth	16	0.88	0.70 – 0.96	$p < 0.001$
Risk assessment for goals	19	0.42	-0.03 – 0.73	$p < 0.05$
Individual & systemic focus	20	0.87	0.70 – 0.95	$p < 0.001$
Connect youth to community	20	0.93	0.83 – 0.97	$p < 0.001$
Involve families	20	0.87	0.69 – 0.94	$p < 0.001$
Prioritize school/job re-entry	20	0.88	0.73 – 0.95	$p < 0.001$
Sufficient total intensity (BP)	20	0.97	0.93 – 0.99	$p < 0.001$
Sufficient monthly contact (ISSP)	20	0.82	0.60 – 0.92	$p < 0.001$
Adequate duration	20	0.75	0.48 – 0.89	$p < 0.001$
Target criminogenic needs	20	0.88	0.73 – 0.95	$p < 0.001$
Target specific needs	20	0.89	0.75 – 0.96	$p < 0.001$
Balance support & supervision	20	0.81	0.58 – 0.92	$p < 0.001$
Proper implementation	20	0.85	0.66 – 0.94	$p < 0.001$
Best practice composite	15	0.92	0.78 – 0.97	$p < 0.001$

Note. ICC = Intra-class Correlation Coefficient; CI = Confidence Interval.

Table 6. Correlations between Recidivism Outcomes and Potential Covariates

	Recidivism outcomes							
Covariates	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Gender	-0.10	0.01	0.07	-0.13	-0.04	-0.14	0.06	-0.06
Age	0.07	-0.02	-0.08	0.07	-0.02	0.18*	0.10	0.05
Aboriginal	-0.07	-0.10	-0.10	-0.13	-0.09	-0.12	0.02	-0.13
No Aboriginal	0.20*	0.17	0.22*	0.23*	0.18*	0.10	-0.10	0.24**
? Aboriginal	-0.13	-0.07	-0.12	-0.10	-0.09	0.01	0.07	-0.11
FASD	0.06	0.15	0.18*	-0.07	-0.01	-0.11	-0.11	-0.01
MH issues	0.19*	0.12	0.10	0.04	0.07	0.04	-0.06	0.10
Low cognitive	0.18*	0.11	0.13	0.05	0.05	-0.05	-0.12	0.07
Addictions	0.12	0.11	0.10	0.04	0.07	0.00	-0.12	0.07
Burnaby	0.12	0.09	0.12	0.06	0.10	-0.04	-0.05	0.13
Victoria	-0.03	0.02	-0.02	0.02	-0.05	0.10	0.00	-0.04
Prince George	-0.14	-0.18*	-0.17	-0.14	-0.09	-0.07	0.09	-0.15
SAVRY	0.28**	0.29**	0.29**	0.20*	0.14	0.12	-0.31**	0.23**

Note. Sample size for 1. n = 158 (missing cases deleted listwise).

Sample size for 2. – 8. n = 126 (missing cases deleted listwise, best practice composite cases only).

1. Any recidivism in the year following ISSP.
2. Number of convictions in the year following ISSP.
3. Number of incidents of offending in the year following ISSP.
4. Number of non-breach convictions in the year following ISSP.
5. Number of violent convictions in the year following ISSP.
6. Number of indictable convictions in the year following ISSP.
7. Time in community in the year following ISSP.
8. Seriousness composite in the year following ISSP.

*significant at $p < 0.05$; **significant at $p < 0.01$.

Table 7. Repeated Measures Analysis of Covariance for Pre- and Post-ISSP Convictions by Risk Level

	Sum of Squares	*df*	Mean square	*F*	*p*
Within-Subjects					
Pre-post ISSP convictions	68.78	1	68.78	16.20	0.00
Pre-post x risk	30.23	1	30.23	7.12	0.01
Pre-post x length	4.58	1	4.58	1.08	0.30
Error	696.13	164	4.24		
Between-subjects					
Risk level	199.77	1	199.77	23.28	0.00
Length of ISSP (days)	38.43	1	38.43	4.49	0.04
Error	1407.06	164	8.58		

Table 8. Repeated Measures Analysis of Covariance for Pre- and Post-ISSP Convictions by Match with ISSP Criteria

	Sum of Squares	*df*	Mean square	*F*	*p*
Within-Subjects					
Pre-post ISSP convictions	104.23	1	104.23	24.42	0.00
Pre-post x match	12.32	1	12.32	2.89	0.09
Pre-post x length	4.30	1	4.30	1.01	0.32
Error	717.16	168	4.27		
Between-subjects					
Match w/ISSP criteria	205.57	1	205.57	24.14	0.00
Length of ISSP (days)	43.45	1	43.45	5.10	0.02
Error	1430.43	168	8.51		

Table 9. Paired *t*-tests for Pre- to Post-ISSP Change in Offending by Risk Level and Fit with ISSP Criteria

	Paired Differences					
	Mean (SD)	**SE mean**	**95% CI**	***t***	***df***	***p***
Pre-post ISSP high risk	2.35 (3.11)	0.26	1.84 – 2.87	9.02	141	0.00
Pre-post ISSP low risk	0.72 (1.27)	0.26	0.19 – 1.25	2.82	24	0.01
Pre-post ISSP criteria match	2.27 (3.15)	0.27	1.74 – 2.80	8.47	137	0.00
Pre-post ISSP no criteria match	1.39 (1.64)	0.28	0.81 – 1.98	4.89	32	0.00

Note. Means reflect the difference scores when the post-ISSP scores were subtracted from the pre-ISSP scores, such that positive values indicate decreases in offending between the pre-ISSP and post-ISSP periods.

Table 10. Frequency of Need Areas Present, Needs Identified on the ISSP Referral, and the Appropriateness of Targets on the Referral

	Rated C or D[a]	Identified on referral	Target appropriate[b]	Over-identified	Under-identified
Need area	*n* (%)	*n* (%)	*n* (%)	*n* (%)	*n* (%)
Family relationships	128 (82.6)	46 (29.7)	59 (38.1)	7 (4.5)	89 (57.4)
Supervision	86 (55.5)	24 (15.5)	83 (53.6)	5 (3.2)	67 (43.2)
Housing	88 (56.8)	12 (7.7)	75 (48.4)	2 (1.3)	78 (50.3)
School/Employment	130 (83.9)	100 (64.5)	95 (61.3)	15 (9.7)	45 (29.0)
Peers	134 (86.4)	58 (37.4)	71 (45.8)	4 (2.6)	80 (51.6)
Substance use	125 (80.6)	79 (51.0)	93 (60.0)	8 (5.2)	54 (34.8)
Recreation	114 (73.5)	67 (43.2)	80 (51.6)	14 (9.0)	61 (39.4)
MH/Behavioural	142 (91.6)	65 (41.9)	66 (42.6)	6 (3.9)	83 (53.5)
Attitudes	115 (74.2)	29 (18.7)	67 (43.3)	1 (0.6)	87 (56.1)

Note. Percentages reflect the total youth for which there was a referral (*n* = 155).

[a] Needs rated C were identified as moderately problematic for youth at the time of ISSP referral; needs rated D were identified as highly problematic.

[b] Targets were considered appropriate where a moderate or high need was identified as a target or where a strength or non-problematic area was not identified as a target.

Table 11. Logistic Regression Analysis: Any Recidivism by the Use of Systematic Risk Assessment to Identify Youth

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.58	0.39	2.17	1	0.14	1.79	0.83 – 3.86
MH issues	0.46	0.52	0.79	1	0.38	1.59	0.57 – 4.41
Low cognitive	0.44	0.46	0.90	1	0.34	1.54	0.63 – 3.80
SAVRY	0.06	0.02	6.30	1	0.01	1.06	1.01 – 1.12
Block 2							
No Aboriginal	0.57	0.40	2.10	1	0.15	1.77	0.82 – 3.84
MH issues	0.49	0.52	0.87	1	0.35	1.63	0.58 – 4.55
Low cognitive	0.41	0.46	0.79	1	0.37	1.51	0.61 – 3.73
SAVRY	0.06	0.02	6.58	1	0.01	1.07	1.02 – 1.12
Best practice 2	-0.10	0.15	0.43	1	0.51	0.90	0.66 – 1.23

Note. Sample size $n = 135$. Model -2 Log likelihood = 169.65, $\chi^2(5) = 14.82$, $p < 0.05$, Cox and Snell $R^2 = 0.10$, Nagelkerke $R^2 = 0.14$, Hosmer-Lemeshow $\chi^2(8) = 4.04$, $p = 0.85$; Block 2 $\chi^2(1) = 0.43$, *ns*.

Table 12. Logistic Regression Analysis: Any Recidivism by the Use of Systematic Risk Assessment to Identify Goals

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.60	0.39	2.38	1	0.12	1.82	0.85 – 3.87
MH issues	0.53	0.52	1.06	1	0.30	1.70	0.62 – 4.68
Low cognitive	0.58	0.45	1.67	1	0.20	1.80	0.74 – 4.34
SAVRY	0.07	0.02	7.07	1	0.01	1.07	1.02 – 1.12
Block 2							
No Aboriginal	0.56	0.39	2.12	1	0.15	1.76	0.82 – 3.77
MH issues	0.55	0.52	1.15	1	0.28	1.74	0.63 – 4.80
Low cognitive	0.58	0.45	1.65	1	0.20	1.78	0.73 – 4.30
SAVRY	0.07	0.03	7.42	1	0.01	1.07	1.02 – 1.13
Best practice 3	0.15	0.20	0.55	1	0.46	1.16	0.79 – 1.70

Note. Sample size $n = 144$. Model -2 Log likelihood = 176.56, $\chi^2(5) = 19.05$, $p < 0.01$, Cox and Snell $R^2 = 0.12$, Nagelkerke $R^2 = 0.17$, Hosmer-Lemeshow $\chi^2(8) = 7.06$, $p = 0.53$; Block 2 $\chi^2(1) = 0.55$, *ns*.

Table 13. Correlations between Contextual Variables

	A	B	C	D
Target individual and systemic factors (A)				
Community involvement (B)	0.84***			
Make contact with families (C)	0.87***	0.71***		
Prioritize school/employment re-entry (D)	0.50***	0.58***	0.39***	

Note. Sample size $n = 175$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 14. Correlations between School/Employment Outcomes and Potential Covariates

	School and employment outcomes							
Covariates	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Gender	-0.27**	-0.28**	-0.33**	-0.13	-0.04	0.16	0.07	0.10
Age	0.07	0.28**	0.20**	0.15	-0.19*	-0.30**	-0.26**	-0.12
Aboriginal	-0.03	-0.10	-0.18*	-0.09	-0.17*	-0.03	0.00	0.09
No Aboriginal	0.00	-0.04	0.05	-0.14	0.07	-0.15	-0.15	0.04
? Aboriginal	0.02	0.12	0.10	0.21*	0.08	0.20*	0.14	-0.12
FASD	-0.16	-0.21*	-0.16	-0.13	0.00	0.08	0.05	0.13
MH issues	-0.03	-0.18*	-0.13	-0.18*	0.05	-0.08	-0.16*	0.09
Low cognitive	-0.08	-0.16	-0.17*	-0.15	-0.08	0.05	-0.02	0.10
Addictions	-0.06	-0.08	-0.12	-0.18*	0.05	-0.01	-0.05	0.22**
Burnaby	0.07	0.04	-0.03	-0.06	0.09	-0.01	-0.04	-0.14
Victoria	-0.08	0.05	0.08	0.11	-0.19*	-0.14	-0.07	0.04
Prince George	0.01	-0.14	-0.06	-0.07	0.12	0.22*	0.17*	0.15
SAVRY	-0.08	0.19*	-0.26**	-0.26**	-0.13*	-0.28**	-0.25**	0.24**
File quality	0.13	0.08	0.00	-0.13	0.13	0.07	0.05	0.33**
File time	0.33**	0.18*	0.22**	0.18*	0.20**	0.16	0.08	0.23**
School start	N/A	N/A	N/A	N/A	0.41**	0.55**	0.28**	N/A
Employ start	0.18*	0.40**	0.45**	0.33**	N/A	N/A	N/A	N/A

Note. Sample size n = 143 (missing data deleted listwise).

1. Any part-time employment during ISSP and follow up.
2. Any full-time employment during ISSP and follow up.
3. Overall maintenance of employment during ISSP and follow up.
4. Employment at the end of follow up.
5. Any school during ISSP or follow up.
6. Overall enrollment in school during ISSP or follow up.
7. School at the end of follow up.
8. Number of community programs in addition to ISSP.

*significant at $p < 0.05$; **significant at $p < 0.01$.

Table 15. Logistic Regression Analysis: Any Part-time Employment by Targeting Individual and Systemic Factors

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Gender	-1.51	0.41	13.75	1	0.00	0.22	0.10 – 0.49
File time (months)	0.11	0.03	16.12	1	0.00	1.12	1.06 – 1.18
Block 2							
Gender	-1.57	0.41	14.28	1	0.00	0.21	0.09 – 0.47
File time (months)	0.10	0.03	12.09	1	0.00	1.10	1.04 – 1.17
Best practice 4	0.28	0.21	1.70	1	0.19	1.32	0.87 – 1.99
Block 3							
Gender	-1.54	0.41	13.92	1	0.00	0.21	0.09 – 0.48
File time (months)	0.16	0.07	5.67	1	0.02	1.18	1.03 – 1.35
Best practice 4	0.70	0.46	2.32	1	0.13	2.01	0.82 – 4.93
FT x BP 4	-0.03	0.03	1.15	1	0.28	0.97	0.92 – 1.02

Note. Sample size $n = 175$. Model -2 Log likelihood = 176.38, $\chi^2(4) = 37.54$, $p < 0.001$, Cox and Snell $R^2 = 0.19$, Nagelkerke $R^2 = 0.27$, Hosmer-Lemeshow $\chi^2(8) = 8.61$, $p = 0.38$; Block 2 $\chi^2(1) = 1.75$, *ns*; Block 3 $\chi^2(1) = 1.12$, *ns*.

Table 16. Logistic Regression Analysis: Any Part-time Employment by Community Involvement

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(B)
Block 1							
Gender	-1.51	0.41	13.75	1	0.00	0.22	0.10 – 0.49
File time (months)	0.11	0.03	16.12	1	0.00	1.12	1.06 – 1.18
Block 2							
Gender	-1.62	0.42	14.92	1	0.00	0.10	0.09 – 0.45
File time (months)	0.09	0.03	10.50	1	0.00	1.10	1.04 – 1.16
Best practice 5	0.36	0.18	3.93	1	0.05	1.43	1.00 – 2.04
Block 3							
Gender	-1.62	0.42	14.98	1	0.00	0.20	0.09 – 0.45
File time (months)	0.12	0.06	3.44	1	0.06	1.13	0.99 – 1.28
Best practice 5	0.52	0.39	1.78	1	0.18	1.68	0.78 – 3.60
FT x BP 5	-0.01	0.02	0.22	1	0.64	0.99	0.94 – 1.04

Note. Sample size $n = 175$. Model -2 Log likelihood = 174.92, $\chi^2(4) = 39.00$, $p < 0.001$, Cox and Snell $R^2 = 0.20$, Nagelkerke $R^2 = 0.28$, Hosmer-Lemeshow $\chi^2(8) = 11.96$, $p = 0.15$; Block 2 $\chi^2(1) = 4.12$, $p < 0.05$; Block 3 $\chi^2(1) = 0.22$, *ns*.

Table 17. Logistic Regression Analysis: Any Part-time Employment by Family Contact

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Gender	-1.51	0.41	13.75	1	0.00	0.22	0.10 – 0.49
File time (months)	0.11	0.03	16.12	1	0.00	1.12	1.06 – 1.18
Block 2							
Gender	-1.51	0.41	13.68	1	0.00	0.22	0.10 – 0.49
File time (months)	0.11	0.03	14.68	1	0.00	1.12	1.06 – 1.18
Best practice 6	0.01	0.17	0.00	1	0.95	1.01	0.72 – 1.42
Block 3							
Gender	-1.51	0.41	13.67	1	0.00	0.22	0.10 – 0.49
File time (months)	0.15	0.06	6.44	1	0.01	1.16	1.03 – 1.31
Best practice 6	0.29	0.38	5.70	1	0.45	1.33	0.63 – 2.82
FT x BP 6	-0.02	0.02	0.68	1	0.41	0.98	0.94 – 1.03

Note. Sample size $n = 175$. Model -2 Log likelihood = 178.58, $\chi^2(4) = 35.33$, $p < 0.001$, Cox and Snell $R^2 = 0.18$, Nagelkerke $R^2 = 0.26$, Hosmer-Lemeshow $\chi^2(8) = 7.95$, $p = 0.44$; Block 2 $\chi^2(1) = 0.00$, *ns*; Block 3 $\chi^2(1) = 0.67$, *ns*.

Table 18. Logistic Regression Analysis: Any Part-time Employment by Prioritizing School/Employment Re-entry

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Gender	-1.51	0.41	13.75	1	0.00	0.22	0.10 – 0.49
File time (months)	0.11	0.03	16.12	1	0.00	1.12	1.06 – 1.18
Block 2							
Gender	-1.50	0.41	13.54	1	0.00	0.22	0.10 – 0.50
File time (months)	0.10	0.03	10.64	1	0.00	1.10	1.04 – 1.16
Best practice 7	0.29	0.20	2.12	1	0.14	1.33	0.91 – 1.96
Block 3							
Gender	-1.53	0.41	13.87	1	0.00	0.22	0.10 – 0.48
File time (months)	0.13	0.06	4.97	1	0.03	1.14	1.02 – 1.29
Best practice 7	0.60	0.46	1.71	1	0.19	1.83	0.74 – 4.53
FT x BP 7	-0.02	0.03	0.59	1	0.44	0.98	0.93 – 1.03

Note. Sample size $n = 175$. Model -2 Log likelihood = 176.46, $\chi^2(4) = 37.46$, $p < 0.001$, Cox and Snell $R^2 = 0.19$, Nagelkerke $R^2 = 0.27$, Hosmer-Lemeshow $\chi^2(8) = 14.66$, $p = 0.07$; Block 2 $\chi^2(1) = 2.22$, *ns*; Block 3 $\chi^2(1) = 0.58$, *ns*.

Table 19. Logistic Regression Analysis: Any Full-time Employment by Targeting Individual and Systemic Factors

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Gender	-1.66	0.67	6.13	1	0.01	0.19	0.05 – 0.71
Age (years)	0.63	0.19	10.50	1	0.00	1.87	1.28 – 2.73
FASD	-0.32	0.41	0.61	1	0.44	0.73	0.32 – 1.62
MH issues	-1.30	0.58	5.08	1	0.02	0.27	0.09 – 0.84
SAVRY	-0.01	0.02	0.06	1	0.81	0.99	0.95 – 1.04
File time (months)	0.08	0.03	8.56	1	0.00	1.08	1.03 – 1.15
Block 2							
Gender	-1.66	0.67	6.14	1	0.01	0.19	0.05 – 0.71
Age (years)	0.61	0.20	9.59	1	0.00	1.84	1.25 – 2.70
FASD	-0.32	0.41	0.59	1	0.44	0.73	0.32 – 1.64
MH issues	-1.31	0.58	5.11	1	0.02	0.27	0.09 – 0.84
SAVRY	0.00	0.02	0.04	1	0.83	1.00	0.95 – 1.04
File time (months)	0.08	0.03	8.69	1	0.00	1.09	1.03 – 1.15
Best practice 4	-0.12	0.24	0.24	1	0.62	0.89	0.55 – 1.43
Block 3							
Gender	-1.68	0.68	6.04	1	0.01	0.19	0.05 – 0.71
Age (years)	0.60	0.20	8.51	1	0.00	1.82	1.22 – 2.71
FASD	-0.28	0.43	0.42	1	0.52	0.76	0.33 – 1.75
MH issues	-1.26	0.58	4.67	1	0.03	0.28	0.09 – 0.89
SAVRY	-0.01	0.02	0.07	1	0.79	0.99	0.95 – 1.04
File time (months)	0.04	0.07	0.23	1	0.63	1.04	0.90 – 1.19
Best practice 4	-0.59	0.62	0.91	1	0.34	0.55	0.16 – 1.87
Age x BP 4	0.38	0.24	2.41	1	0.12	1.46	0.90 – 2.36
FT x BP 4	0.02	0.03	0.51	1	0.48	1.04	0.96 – 1.09

Note. Sample size $n = 156$. Model -2 Log likelihood = 156.08, $\chi^2(9) = 39.63$, $p < 0.001$, Cox and Snell $R^2 = 0.22$, Nagelkerke $R^2 = 0.31$, Hosmer-Lemeshow $\chi^2(8) = 11.64$, $p = 0.17$; Block 2 $\chi^2(1) = 0.24$, *ns*; Block 3 $\chi^2(2) = 2.73$, *ns*.

Table 20. Logistic Regression Analysis: Any Full-time Employment by Community Involvement

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Gender	-1.66	0.67	6.13	1	0.01	0.19	0.05 – 0.71
Age (years)	0.63	0.19	10.50	1	0.00	1.87	1.28 – 2.73
FASD	-0.32	0.41	0.61	1	0.44	0.73	0.32 – 1.62
MH issues	-1.30	0.58	5.08	1	0.02	0.27	0.09 – 0.84
SAVRY	-0.01	0.02	0.06	1	0.81	0.99	0.95 – 1.04
File time (months)	0.08	0.03	8.56	1	0.00	1.08	1.03 – 1.15
Block 2							
Gender	-1.66	0.67	6.20	1	0.01	0.19	0.05 – 0.70
Age (years)	0.58	0.20	8.78	1	0.00	1.79	1.22 – 2.64
FASD	-0.33	0.42	0.61	1	0.42	0.72	0.32 – 1.62
MH issues	-1.35	0.58	5.32	1	0.02	0.26	0.08 – 0.82
SAVRY	0.00	0.02	0.01	1	0.92	1.00	0.95 – 1.05
File time (months)	0.09	0.03	9.48	1	0.00	1.10	1.03 – 1.16
Best practice 5	-0.22	0.20	1.10	1	0.29	0.81	0.54 – 1.20
Block 3							
Gender	-1.71	0.69	6.35	1	0.01	0.18	0.05 – 0.68
Age (years)	0.58	0.20	8.41	1	0.00	1.80	1.21 – 2.67
FASD	-0.35	0.42	0.69	1	0.41	0.70	0.31 – 1.62
MH issues	-1.27	0.59	4.73	1	0.03	0.28	0.09 – 0.88
SAVRY	-0.01	0.02	0.07	1	0.80	0.99	0.95 – 1.04
File time (months)	0.03	0.07	0.25	1	0.62	1.03	0.91 – 1.18
Best practice 5	-0.68	0.50	1.86	1	0.17	0.50	0.19 – 1.35
Age x BP 5	0.20	0.19	1.04	1	0.31	1.22	0.83 – 1.78
FT x BP 5	0.03	0.03	0.93	1	0.33	1.03	0.97 – 1.09

Note. Sample size $n = 156$. Model -2 Log likelihood = 156.47, $\chi^2(9) = 39.23$, $p < 0.001$, Cox and Snell $R^2 = 0.22$, Nagelkerke $R^2 = 0.31$, Hosmer-Lemeshow $\chi^2(8) = 10.38$, $p = 0.24$; Block 2 $\chi^2(1) = 1.12$, *ns*; Block 3 $\chi^2(2) = 1.46$, *ns*.

Table 21. Logistic Regression Analysis: Any Full-time Employment by Family Involvement

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Gender	-1.66	0.67	6.13	1	0.01	0.19	0.05 – 0.71
Age (years)	0.63	0.19	10.50	1	0.00	1.87	1.28 – 2.73
FASD	-0.32	0.41	0.61	1	0.44	0.73	0.32 – 1.62
MH issues	-1.30	0.58	5.08	1	0.02	0.27	0.09 – 0.84
SAVRY	-0.01	0.02	0.06	1	0.81	0.99	0.95 – 1.04
File time (months)	0.08	0.03	8.56	1	0.00	1.08	1.03 – 1.15
Block 2							
Gender	-1.65	0.67	6.03	1	0.01	0.19	0.05 – 0.72
Age (years)	0.56	0.20	7.72	1	0.00	1.75	1.18 – 2.60
FASD	-0.32	0.42	0.56	1	0.46	0.73	0.32 – 1.66
MH issues	-1.24	0.58	4.57	1	0.03	0.29	0.09 – 0.90
SAVRY	-0.01	0.02	0.09	1	0.77	0.99	0.95 – 1.04
File time (months)	0.09	0.03	9.97	1	0.00	1.10	1.04 – 1.16
Best practice 6	-0.29	0.21	1.85	1	0.17	0.75	0.49 – 1.14
Block 3							
Gender	-1.66	0.68	5.97	1	0.02	0.19	0.05 – 0.72
Age (years)	0.55	0.21	6.91	1	0.01	1.74	1.15 – 2.62
FASD	-0.28	0.43	0.40	1	0.52	0.76	0.32 – 1.78
MH issues	-1.28	0.59	4.77	1	0.03	0.28	0.09 – 0.88
SAVRY	-0.01	0.02	0.13	1	0.72	0.99	0.94 – 1.04
File time (months)	0.05	0.06	0.79	1	0.37	1.06	0.94 – 1.19
Best practice 6	-0.74	0.58	1.63	1	0.20	0.48	0.16 – 1.48
Age x BP 6	0.30	0.20	2.27	1	0.13	1.36	0.91 – 2.01
FT x BP 6	0.02	0.03	0.50	1	0.48	1.02	0.96 – 1.08

Note. Sample size n = 156. Model -2 Log likelihood = 154.54, $\chi^2(9) = 41.16$, $p < 0.001$, Cox and Snell $R^2 = 0.23$, Nagelkerke $R^2 = 0.32$, Hosmer-Lemeshow $\chi^2(8) = 9.12$, $p = 0.33$; Block 2 $\chi^2(1) = 1.90$, *ns*; Block 3 $\chi^2(2) = 2.61$, *ns*.

Table 22. Logistic Regression Analysis: Any Full-time Employment by Prioritizing School/Employment Re-entry

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Gender	-1.66	0.67	6.13	1	0.01	0.19	0.05 – 0.71
Age (years)	0.63	0.19	10.50	1	0.00	1.87	1.28 – 2.73
FASD	-0.32	0.41	0.61	1	0.44	0.73	0.32 – 1.62
MH issues	-1.30	0.58	5.08	1	0.02	0.27	0.09 – 0.84
SAVRY	-0.01	0.02	0.06	1	0.81	0.99	0.95 – 1.04
File time (months)	0.08	0.03	8.56	1	0.00	1.08	1.03 – 1.15
Block 2							
Gender	-1.71	0.68	6.38	1	0.01	0.18	0.05 – 0.68
Age (years)	0.63	0.19	10.69	1	0.00	1.88	1.29 – 2.75
FASD	-0.31	0.41	0.58	1	0.48	0.73	0.33 – 1.63
MH issues	-1.36	0.58	5.41	1	0.02	0.26	0.08 – 0.81
SAVRY	0.00	0.02	0.01	1	0.93	1.00	0.95 – 1.05
File time (months)	0.10	0.03	9.54	1	0.00	1.10	1.04 – 1.17
Best practice 7	-0.22	0.20	1.27	1	0.26	0.80	0.55 – 1.18
Block 3							
Gender	-1.72	0.68	6.38	1	0.01	0.18	0.05 – 0.68
Age (years)	0.64	0.19	10.66	1	0.00	1.89	1.29 – 2.77
FASD	-0.31	0.41	0.55	1	0.46	0.74	0.33 – 1.65
MH issues	-1.36	0.58	5.41	1	0.02	0.26	0.08 – 0.81
SAVRY	0.00	0.02	0.00	1	0.94	1.00	0.95 – 1.05
File time (months)	0.10	0.06	3.02	1	0.08	1.11	0.99 – 1.25
Best practice 7	-0.14	0.51	0.08	1	0.78	0.87	0.32 – 2.36
FT x BP 7	0.00	0.03	0.03	1	0.87	1.00	0.94 – 1.05

Note. Sample size $n = 156$. Model -2 Log likelihood = 157.73, $\chi^2(8) = 37.98$, $p < 0.001$, Cox and Snell $R^2 = 0.22$, Nagelkerke $R^2 = 0.30$, Hosmer-Lemeshow $\chi^2(8) = 3.84$, $p = 0.87$; Block 2 $\chi^2(1) = 1.29$, *ns*; Block 3 $\chi^2(1) = 0.03$, *ns*.

Table 23. Linear Regression Analysis: Overall Employment by Targeting Individual and Systemic Factors

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.36	$F(7, 150) = 12.26^{***}$
Gender	-0.35	0.11	-0.22	-3.11	0.00		
Age (years)	0.06	0.04	0.12	1.71	0.09		
Aboriginal	-0.10	0.11	-0.06	-0.91	0.36		
Low cognitive	-0.13	0.11	-0.08	-1.14	0.26		
SAVRY	-0.01	0.00	-0.10	-1.47	0.14		
File time (months)	0.02	0.01	0.19	2.88	0.00		
Employment start	0.71	0.13	0.38	5.62	0.00		
Block 2						0.00	$F(1, 149) = 0.03$
Gender	-0.35	0.11	-0.21	-3.10	0.00		
Age (years)	0.06	0.04	0.11	1.55	0.12		
Aboriginal	-0.10	0.11	-0.06	-0.89	0.38		
Low cognitive	-0.13	0.11	-0.08	-1.12	0.26		
SAVRY	-0.01	0.00	-0.10	-1.47	0.14		
File time (months)	0.02	0.01	0.20	2.76	0.01		
Employment start	0.71	0.13	0.38	5.58	0.00		
Best practice 4	-0.01	0.05	-0.01	-0.17	0.86		
Block 3						0.00	$F(2, 147) = 0.26$
Gender	-0.34	0.11	-0.21	-3.00	0.00		
Age (years)	0.06	0.04	0.11	1.46	0.15		
Aboriginal	-0.09	0.11	-0.06	-0.85	0.40		
Low cognitive	-0.12	0.12	-0.07	-1.00	0.32		
SAVRY	-0.01	0.01	-0.10	-1.46	0.15		
File time (months)	0.03	0.02	0.30	1.78	0.08		
Employment start	0.71	0.13	0.38	5.53	0.00		
Best practice 4	0.06	0.12	0.08	0.52	0.60		
Age x BP 4	0.00	0.04	0.00	0.01	0.99		
FT x BP 4	0.00	0.01	-0.17	-0.67	0.50		

Note. Sample size $n = 158$.

*significant at $p < 0.05$, **significant at $p < 0.01$, ***significant at $p < 0.001$.

Table 24. Linear Regression Analysis: Overall Employment by Community Involvement

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.36	*F*(7, 150) = 12.26***
Gender	-0.35	0.11	-0.22	-3.11	0.00		
Age (years)	0.06	0.04	0.12	1.71	0.09		
Aboriginal	-0.10	0.11	-0.06	-0.91	0.36		
Low cognitive	-0.13	0.11	-0.08	-1.14	0.26		
SAVRY	-0.01	0.00	-0.10	-1.47	0.14		
File time (months)	0.02	0.01	0.19	2.88	0.00		
Employment start	0.71	0.13	0.38	5.62	0.00		
Block 2						0.00	*F*(1, 149) = 0.15
Gender	-0.35	0.11	-0.22	-3.11	0.00		
Age (years)	0.06	0.04	0.11	1.55	0.12		
Aboriginal	-0.10	0.11	-0.06	-0.88	0.38		
Low cognitive	-0.13	0.11	-0.08	-1.12	0.26		
SAVRY	-0.01	0.00	-0.10	-1.45	0.15		
File time (months)	0.02	0.01	0.20	2.80	0.01		
Employment start	0.71	0.13	0.38	5.60	0.00		
Best practice 5	-0.02	0.05	-0.03	-0.38	0.70		
Block 3						0.00	*F*(2, 147) = 0.08
Gender	-0.35	0.11	-0.22	-3.09	0.00		
Age (years)	0.06	0.04	0.11	1.46	0.15		
Aboriginal	-0.09	0.11	-0.06	-0.85	0.40		
Low cognitive	-0.12	0.12	-0.08	-1.07	0.29		
SAVRY	-0.01	0.01	-0.10	-1.42	0.16		
File time (months)	0.02	0.02	0.26	1.58	0.12		
Employment start	0.70	0.13	0.38	5.51	0.00		
Best practice 5	0.02	0.10	0.03	0.18	0.86		
Age x BP 5	0.00	0.04	0.00	-0.07	0.94		
FT x BP 5	0.00	0.01	-0.10	-0.40	0.69		

Note. Sample size $n = 158$.

*significant at $p < 0.05$, **significant at $p < 0.01$, ***significant at $p < 0.001$.

Table 25. Linear Regression Analysis: Overall Employment by Family Contact

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.36	*F*(7, 150) = 12.26***
Gender	-0.35	0.11	-0.22	-3.11	0.00		
Age (years)	0.06	0.04	0.12	1.71	0.09		
Aboriginal	-0.10	0.11	-0.06	-0.91	0.36		
Low cognitive	-0.13	0.11	-0.08	-1.14	0.26		
SAVRY	-0.01	0.00	-0.10	-1.47	0.14		
File time (months)	0.02	0.01	0.19	2.88	0.00		
Employment start	0.71	0.13	0.38	5.62	0.00		
Block 2						0.01	*F*(1, 149) = 1.87
Gender	-0.35	0.11	-0.21	-3.08	0.00		
Age (years)	0.04	0.04	0.08	1.03	0.30		
Aboriginal	-0.09	0.11	-0.06	-0.81	0.42		
Low cognitive	-0.11	0.11	-0.06	-0.94	0.35		
SAVRY	-0.01	0.00	-0.11	-1.63	0.10		
File time (months)	0.02	0.01	0.22	3.15	0.00		
Employment start	0.72	0.13	0.39	5.74	0.00		
Best practice 6	-0.06	0.05	-0.10	-1.37	0.17		
Block 3						0.01	*F*(2, 147) = 0.78
Gender	-0.34	0.11	-0.22	-3.02	0.00		
Age (years)	0.04	0.04	0.11	0.84	0.40		
Aboriginal	-0.08	0.11	-0.06	-0.70	0.49		
Low cognitive	-0.09	0.11	-0.08	-0.82	0.41		
SAVRY	-0.01	0.00	-0.10	-1.59	0.11		
File time (months)	0.03	0.01	0.26	2.53	0.01		
Employment start	0.72	0.13	0.38	5.73	0.00		
Best practice 6	0.04	0.10	0.03	0.35	0.72		
Age x BP 6	0.01	0.03	0.00	0.28	0.78		
FT x BP 6	-0.01	0.00	-0.10	-1.11	0.27		

Note. Sample size $n = 158$.

*significant at $p < 0.05$, **significant at $p < 0.01$, ***significant at $p < 0.001$.

Table 26. Linear Regression Analysis: Overall Employment by Prioritizing School/Employment Re-entry

Predictor	*B*	SE(B)	*B*	*t*	*p*	Δ R^2	*F* change
Block 1						0.36	*F*(7, 150) = 12.26***
Gender	-0.35	0.11	-0.22	-3.11	0.00		
Age (years)	0.06	0.04	0.12	1.71	0.09		
Aboriginal	-0.10	0.11	-0.06	-0.91	0.36		
Low cognitive	-0.13	0.11	-0.08	-1.14	0.26		
SAVRY	-0.01	0.00	-0.10	-1.47	0.14		
File time (months)	0.02	0.01	0.19	2.88	0.00		
Employment start	0.71	0.13	0.38	5.62	0.00		
Block 2						0.00	*F*(1, 149) = 0.07
Gender	-0.35	0.11	-0.21	-3.07	0.00		
Age (years)	0.06	0.04	0.12	1.71	0.09		
Aboriginal	-0.10	0.11	-0.06	-0.92	0.36		
Low cognitive	-0.13	0.11	-0.08	-1.13	0.26		
SAVRY	-0.01	0.00	-0.10	-1.48	0.14		
File time (months)	0.02	0.01	0.18	2.46	0.01		
Employment start	0.71	0.13	0.39	5.60	0.00		
Best practice 7	0.01	0.04	-0.02	0.27	0.79		
Block 3						0.00	*F*(1, 148) = 0.01
Gender	-0.34	0.11	-0.21	-3.05	0.00		
Age (years)	0.06	0.04	0.12	1.69	0.09		
Aboriginal	-0.10	0.11	-0.06	-0.91	0.36		
Low cognitive	-0.13	0.11	-0.08	-1.13	0.26		
SAVRY	-0.01	0.01	-0.10	-1.48	0.14		
File time (months)	0.02	0.01	0.17	1.18	0.24		
Employment start	0.71	0.13	0.38	5.58	0.00		
Best practice 7	0.00	0.11	0.00	0.02	0.98		
FT x BP 7	0.00	0.01	0.02	0.10	0.92		

Note. Sample size *n* = 158.

*significant at $p < 0.05$, **significant at $p < 0.01$, ***significant at $p < 0.001$.

Table 27. Logistic Regression Analysis: Employment at the End of Follow up by Targeting Individual and Systemic Factors

Predictor	B	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
? Aboriginal	0.57	0.42	1.86	1	0.17	1.77	0.78 – 4.00
MH issues	-1.24	0.76	2.65	1	0.10	0.29	0.06 – 1.29
Addictions	-0.51	0.45	1.30	1	0.25	0.60	0.25 – 1.44
SAVRY	-0.04	0.03	2.47	1	0.12	0.96	0.91 – 1.01
File time (months)	0.07	0.03	6.06	1	0.01	1.07	1.02 – 1.14
Employment start	1.84	0.54	11.65	1	0.00	6.32	2.19 – 18.22
Block 2							
? Aboriginal	0.57	0.42	1.86	1	0.17	1.77	0.78 – 4.00
MH issues	-1.24	0.76	2.66	1	0.10	0.29	0.06 – 1.29
Addictions	-0.51	0.45	1.30	1	0.25	0.60	0.25 – 1.44
SAVRY	-0.04	0.03	2.46	1	0.12	0.96	0.91 – 1.01
File time (months)	0.07	0.03	5.43	1	0.02	1.07	1.01 – 1.14
Employment start	1.84	0.54	11.63	1	0.00	6.32	2.19 – 18.22
Best practice 4	-0.01	0.23	0.00	1	0.96	0.99	0.63 – 1.56
Block 3							
? Aboriginal	0.56	0.42	1.78	1	0.18	1.75	0.77 – 3.97
MH issues	-1.25	0.76	2.68	1	0.10	0.29	0.06 – 1.28
Addictions	-0.56	0.46	1.51	1	0.22	0.57	0.23 – 1.40
SAVRY	-0.04	0.03	2.30	1	0.13	0.96	0.91 – 1.01
File time (months)	0.12	0.08	2.30	1	0.13	1.13	0.97 – 1.31
Employment start	1.86	0.54	11.68	1	0.00	6.43	2.21 – 18.71
Best practice 4	0.37	0.62	0.34	1	0.56	1.44	0.42 – 4.90
FT x BP 4	-0.02	0.03	0.42	1	0.52	0.98	0.92 – 1.04

Note. Sample size n = 152. Model -2 Log likelihood = 144.78, $\chi^2(8) = 36.30$, $p < 0.001$, Cox and Snell R^2 = 0.21, Nagelkerke R^2 = 0.30, Hosmer-Lemeshow $\chi^2(8) = 12.48$, $p = 0.13$; Block 2 $\chi^2(1) = 0.00$, *ns*; Block 3 $\chi^2(1) = 0.44$, *ns*.

Table 28. Logistic Regression Analysis: Employment at the End of Follow up by Community Involvement

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
? Aboriginal	0.57	0.42	1.86	1	0.17	1.77	0.78 – 4.00
MH issues	-1.24	0.76	2.65	1	0.10	0.29	0.06 – 1.29
Addictions	-0.51	0.45	1.30	1	0.25	0.60	0.25 – 1.44
SAVRY	-0.04	0.03	2.47	1	0.12	0.96	0.91 – 1.01
File time (months)	0.07	0.03	6.06	1	0.01	1.07	1.02 – 1.14
Employment start	1.84	0.54	11.65	1	0.00	6.32	2.19 – 18.22
Block 2							
? Aboriginal	0.56	0.42	1.80	1	0.18	1.75	0.77 – 3.97
MH issues	-1.24	0.77	2.59	1	0.11	0.29	0.06 – 1.31
Addictions	-0.51	0.45	1.28	1	0.26	0.60	0.25 – 1.45
SAVRY	-0.04	0.03	2.54	1	0.11	0.96	0.91 – 1.01
File time (months)	0.07	0.03	4.63	1	0.03	1.07	1.01 – 1.14
Employment start	1.88	0.55	11.80	1	0.00	6.54	2.24 – 19.08
Best practice 5	0.08	0.21	0.15	1	0.70	1.08	0.72 – 1.63
Block 3							
? Aboriginal	0.53	0.42	1.59	1	0.21	1.70	0.74 – 3.89
MH issues	-1.24	0.77	2.62	1	0.10	0.29	0.06 – 1.30
Addictions	-0.53	0.45	1.38	1	0.24	0.59	0.24 – 1.43
SAVRY	-0.04	0.03	2.50	1	0.11	0.96	0.91 – 1.01
File time (months)	0.11	0.08	1.99	1	0.16	1.11	0.96 – 1.30
Employment start	1.89	0.55	11.72	1	0.00	6.62	2.24 – 19.54
Best practice 5	0.36	0.52	0.48	1	0.49	1.43	0.52 – 3.95
FT x BP 5	-0.02	0.03	0.34	1	0.56	0.98	0.93 – 1.04

Note. Sample size n = 152. Model -2 Log likelihood = 144.73, $\chi^2(8) = 36.36$, $p < 0.001$, Cox and Snell R^2 = 0.21, Nagelkerke R^2 = 0.31, Hosmer-Lemeshow $\chi^2(8) = 13.08$, $p = 0.11$; Block 2 $\chi^2(1) = 0.15$, *ns*; Block 3 $\chi^2(1) = 0.35$, *ns*.

Table 29. Logistic Regression Analysis: Employment at the End of Follow up by Family Contact

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
? Aboriginal	0.57	0.42	1.86	1	0.17	1.77	0.78 – 4.00
MH issues	-1.24	0.76	2.65	1	0.10	0.29	0.06 – 1.29
Addictions	-0.51	0.45	1.30	1	0.25	0.60	0.25 – 1.44
SAVRY	-0.04	0.03	2.47	1	0.12	0.96	0.91 – 1.01
File time (months)	0.07	0.03	6.06	1	0.01	1.07	1.02 – 1.14
Employment start	1.84	0.54	11.65	1	0.00	6.32	2.19 – 18.22
Block 2							
? Aboriginal	0.57	0.42	1.86	1	0.17	1.77	0.78 – 4.04
MH issues	-1.22	0.75	2.65	1	0.10	0.30	0.07 – 1.28
Addictions	-0.51	0.45	1.27	1	0.26	0.60	0.25 – 1.46
SAVRY	-0.04	0.03	2.58	1	0.11	0.96	0.91 – 1.01
File time (months)	0.08	0.03	7.18	1	0.01	1.09	1.02 – 1.16
Employment start	1.81	0.55	10.95	1	0.00	6.14	2.10 – 17.96
Best practice 6	-0.24	0.20	1.36	1	0.24	0.79	0.53 – 1.18
Block 3							
? Aboriginal	0.57	0.42	1.81	1	0.18	1.76	0.77 – 4.02
MH issues	-1.22	0.74	2.68	1	0.10	0.30	0.07 – 1.27
Addictions	-0.56	0.46	1.45	1	0.23	0.57	0.23 – 1.42
SAVRY	-0.04	0.03	2.34	1	0.13	0.96	0.91 – 1.01
File time (months)	0.11	0.07	2.83	1	0.09	1.12	0.98 – 1.28
Employment start	1.83	0.55	11.05	1	0.00	6.24	2.12 – 18.37
Best practice 6	0.03	0.56	0.00	1	0.96	1.03	0.34 – 3.07
FT x BP 6	-0.01	0.03	0.25	1	0.61	0.99	0.93 – 1.04

Note. Sample size n = 152. Model -2 Log likelihood = 143.58, $\chi^2(8) = 37.51$, $p < 0.001$, Cox and Snell R^2 = 0.22, Nagelkerke R^2 = 0.31, Hosmer-Lemeshow $\chi^2(8) = 12.94$, $p = 0.11$; Block 2 $\chi^2(1) = 1.39$, *ns*; Block 3 $\chi^2(1) = 0.26$, *ns*.

Table 30. Logistic Regression Analysis: Employment at the End of Follow up by Prioritizing School/Employment Re-entry

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
? Aboriginal	0.57	0.42	1.86	1	0.17	1.77	0.78 – 4.00
MH issues	-1.24	0.76	2.65	1	0.10	0.29	0.06 – 1.29
Addictions	-0.51	0.45	1.30	1	0.25	0.60	0.25 – 1.44
SAVRY	-0.04	0.03	2.47	1	0.12	0.96	0.91 – 1.01
File time (months)	0.07	0.03	6.06	1	0.01	1.07	1.02 – 1.14
Employment start	1.84	0.54	11.65	1	0.00	6.32	2.19 – 18.22
Block 2							
? Aboriginal	0.49	0.43	1.33	1	0.25	1.63	0.71 – 3.77
MH issues	-1.27	0.79	2.58	1	0.11	0.28	0.06 – 1.32
Addictions	-0.61	0.46	1.77	1	0.18	0.54	0.22 – 1.33
SAVRY	-0.05	0.03	3.02	1	0.08	0.95	0.90 – 1.01
File time (months)	0.05	0.03	2.08	1	0.15	1.05	0.98 – 1.12
Employment start	2.08	0.56	13.61	1	0.00	8.02	2.65 – 24.21
Best practice 7	0.47	0.21	4.82	1	0.03	1.60	1.05 – 2.43
Block 3							
? Aboriginal	0.52	0.43	1.45	1	0.23	1.68	0.72 – 3.94
MH issues	-1.37	0.79	2.98	1	0.08	0.25	0.05 – 1.20
Addictions	-0.57	1.02	0.31	1	0.58	1.77	0.24 – 13.13
SAVRY	-0.05	0.03	3.27	1	0.07	0.95	0.90 – 1.00
File time (months)	0.00	0.07	0.00	1	0.98	1.00	0.88 – 1.14
Employment start	2.14	0.57	14.19	1	0.00	8.49	2.79 – 25.85
Best practice 7	0.28	0.58	0.24	1	0.63	1.32	0.43 – 4.08
Addictions x BP 7	-0.52	0.42	1.58	1	0.21	0.59	0.26 – 1.34
FT x BP 7	0.02	0.03	0.54	1	0.46	1.02	0.96 – 1.09

Note. Sample size n = 152. Model -2 Log likelihood = 138.39, $\chi^2(9) = 42.70$, $p < 0.001$, Cox and Snell R^2 = 0.24, Nagelkerke R^2 = 0.35, Hosmer-Lemeshow $\chi^2(8) = 11.25$, $p = 0.19$; Block 2 $\chi^2(1) = 4.99$, $p < 0.05$; Block 3 $\chi^2(1) = 1.85$, *ns*.

Table 31. Logistic Regression Analysis: Any School by Targeting Individual and Systemic Factors

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Age (years)	-0.60	0.28	4.59	1	0.03	0.55	0.32 – 0.95
Aboriginal	-0.56	0.58	0.93	1	0.33	0.57	0.19 – 1.77
Victoria	-0.84	0.54	2.44	1	0.12	0.43	0.15 – 1.24
SAVRY	-0.07	0.04	3.22	1	0.07	0.93	0.86 – 1.01
File time (months)	0.06	0.04	2.43	1	0.12	1.06	0.98 – 1.14
Block 2							
Age (years)	-0.59	0.28	4.40	1	0.04	0.55	0.32 – 0.96
Aboriginal	-0.56	0.58	0.95	1	0.33	0.57	0.18 – 1.77
Victoria	-0.84	0.54	2.41	1	0.12	0.43	0.15 – 1.25
SAVRY	-0.07	0.04	3.20	1	0.07	0.93	0.86 – 1.01
File time (months)	0.06	0.04	1.99	1	0.16	1.06	0.98 – 1.14
Best practice 4	0.04	0.34	0.02	1	0.90	1.04	0.54 – 2.01
Block 3							
Age (years)	-0.64	0.30	4.42	1	0.04	0.53	0.29 – 0.96
Aboriginal	-0.65	0.59	1.22	1	0.27	0.52	0.16 – 1.66
Victoria	-0.74	0.55	1.82	1	0.18	0.48	0.16 – 1.40
SAVRY	-0.07	0.04	3.05	1	0.08	0.94	0.87 – 1.01
File time (months)	0.06	0.10	0.32	1	0.57	1.06	0.87 – 1.28
Best practice 4	0.22	0.71	0.10	1	0.76	1.25	0.31 – 5.01
Age x BP 4	-0.35	0.36	0.96	1	0.33	0.70	0.35 – 1.42
FT x BP 4	0.00	0.05	0.00	1	0.98	1.00	0.91 – 1.10

Note. Sample size n = 162. Model -2 Log likelihood = 99.12, $\chi^2(8) = 21.98$, $p < 0.01$, Cox and Snell R^2 = 0.13, Nagelkerke R^2 = 0.24, Hosmer-Lemeshow $\chi^2(8) = 3.52$, $p = 0.88$; Block 2 $\chi^2(1) = 0.02$, *ns*; Block 3 $\chi^2(2) = 1.12$, *ns*.

Table 32. Logistic Regression Analysis: Any School by Community Involvement

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Age (years)	-0.60	0.28	4.59	1	0.03	0.55	0.32 – 0.95
Aboriginal	-0.56	0.58	0.93	1	0.33	0.57	0.19 – 1.77
Victoria	-0.84	0.54	2.44	1	0.12	0.43	0.15 – 1.24
SAVRY	-0.07	0.04	3.22	1	0.07	0.93	0.86 – 1.01
File time (months)	0.06	0.04	2.43	1	0.12	1.06	0.98 – 1.14
Block 2							
Age (years)	-0.58	0.28	4.07	1	0.04	0.56	0.32 – 0.98
Aboriginal	-0.62	0.58	1.15	1	0.28	0.54	0.17 – 1.68
Victoria	-0.86	0.54	2.50	1	0.11	0.42	0.14 – 1.23
SAVRY	-0.07	0.04	3.31	1	0.07	0.93	0.86 – 1.01
File time (months)	0.04	0.04	1.02	1	0.31	1.04	0.96 – 1.13
Best practice 5	0.30	0.27	1.24	1	0.27	1.35	0.79 – 2.30
Block 3							
Age (years)	-0.62	0.31	4.02	1	0.04	0.54	0.30 – 0.99
Aboriginal	-0.64	0.59	1.20	1	0.27	0.53	0.17 – 1.66
Victoria	-0.82	0.55	2.23	1	0.14	0.44	0.15 – 1.29
SAVRY	-0.07	0.04	3.16	1	0.08	0.93	0.86 – 1.01
File time (months)	0.03	0.09	0.14	1	0.71	1.03	0.87 – 1.22
Best practice 5	0.34	0.61	0.31	1	0.58	1.41	0.43 – 4.64
Age x BP 5	-0.16	0.30	0.29	1	0.59	0.85	0.47 – 1.53
FT x BP 5	0.00	0.04	0.01	1	0.92	1.00	0.93 – 1.08

Note. Sample size n = 162. Model -2 Log likelihood = 98.61, $\chi^2(8) = 22.49$, $p < 0.01$, Cox and Snell R^2 = 0.13, Nagelkerke R^2 = 0.25, Hosmer-Lemeshow $\chi^2(8) = 8.49$, $p = 0.39$; Block 2 $\chi^2(1) = 1.27$, *ns*; Block 3 $\chi^2(2) = 0.37$, *ns*.

Table 33. Logistic Regression Analysis: Any School by Family Involvement

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Age (years)	-0.60	0.28	4.59	1	0.03	0.55	0.32 – 0.95
Aboriginal	-0.56	0.58	0.93	1	0.33	0.57	0.19 – 1.77
Victoria	-0.84	0.54	2.44	1	0.12	0.43	0.15 – 1.24
SAVRY	-0.07	0.04	3.22	1	0.07	0.93	0.86 – 1.01
File time (months)	0.06	0.04	2.43	1	0.12	1.06	0.98 – 1.14
Block 2							
Age (years)	-0.65	0.28	5.30	1	0.02	0.52	0.30 – 0.91
Aboriginal	-0.54	0.58	0.88	1	0.35	0.58	0.19 – 1.80
Victoria	-0.90	0.55	2.70	1	0.10	0.41	0.14 – 1.19
SAVRY	-0.07	0.04	3.54	1	0.06	0.93	0.86 – 1.00
File time (months)	0.07	0.04	3.02	1	0.08	1.07	0.99 – 1.16
Best practice 6	-0.24	0.28	0.76	1	0.38	0.78	0.45 – 1.36
Block 3							
Age (yrs)	-0.67	0.31	4.55	1	0.03	0.51	0.28 – 0.95
Aboriginal	-0.90	0.62	2.11	1	0.15	0.41	0.12 – 1.37
Victoria	-0.06	1.25	0.00	1	0.96	0.94	0.08 – 10.91
SAVRY	-0.07	0.04	3.30	1	0.07	0.93	0.87 – 1.00
File time (months)	0.07	0.09	0.61	1	0.44	1.07	0.90 – 1.27
Best practice 6	0.17	0.78	0.05	1	0.83	1.19	0.26 – 5.47
Age x BP 6	-0.54	0.34	2.42	1	0.12	0.58	0.30 – 1.15
Victoria x BP 6	-0.43	0.68	0.40	1	0.53	0.65	0.17 – 2.45
FT x BP 6	0.00	0.04	0.00	1	0.98	1.00	0.92 – 1.09

Note. Sample size n = 162. Model -2 Log likelihood = 96.29, $\chi^2(9) = 24.81$, $p < 0.01$, Cox and Snell R^2 = 0.14, Nagelkerke R^2 = 0.27, Hosmer-Lemeshow $\chi^2(8) = 9.41$, $p = 0.31$; Block 2 $\chi^2(1) = 0.76$, *ns*; Block 3 $\chi^2(3) = 3.21$, *ns*.

Table 34. Logistic Regression Analysis: Any School by Prioritizing School/Employment Re-entry

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Age (years)	-0.60	0.28	4.59	1	0.03	0.55	0.32 – 0.95
Aboriginal	-0.56	0.58	0.93	1	0.33	0.57	0.19 – 1.77
Victoria	-0.84	0.54	2.44	1	0.12	0.43	0.15 – 1.24
SAVRY	-0.07	0.04	3.22	1	0.07	0.93	0.86 – 1.01
File time (months)	0.06	0.04	2.43	1	0.12	1.06	0.98 – 1.14
Block 2							
Age (years)	-0.60	0.28	4.48	1	0.03	0.55	0.31 – 0.96
Aboriginal	-0.60	0.58	1.08	1	0.30	0.55	0.18 – 1.70
Victoria	-0.94	0.55	2.88	1	0.09	0.39	0.13 – 1.16
SAVRY	-0.07	0.04	3.28	1	0.07	0.93	0.86 – 1.01
File time (months)	0.04	0.04	0.80	1	0.37	1.04	0.96 – 1.12
Best practice 7	0.32	0.29	1.24	1	0.26	1.38	0.78 – 1.12
Block 3							
Age (years)	-0.62	0.29	4.46	1	0.04	0.54	0.30 – 0.96
Aboriginal	-0.67	0.59	1.28	1	0.26	0.51	0.16 – 1.63
Victoria	-0.95	0.56	2.90	1	0.09	0.38	0.13 – 1.16
SAVRY	-0.07	0.04	3.24	1	0.07	0.93	0.86 – 1.01
File time (months)	0.11	0.07	2.25	1	0.13	1.11	0.97 – 1.28
Best practice 7	0.92	0.58	2.50	1	0.11	2.50	0.80 – 7.82
FT x BP 7	-0.04	0.03	1.54	1	0.22	0.96	0.90 – 1.02

Note. Sample size $n = 162$. Model -2 Log likelihood = 97.51, $\chi^2(7) = 23.59$, $p < 0.01$, Cox and Snell $R^2 = 0.14$, Nagelkerke $R^2 = 0.26$, Hosmer-Lemeshow $\chi^2(8) = 6.22$, $p = 0.62$; Block 2 $\chi^2(1) = 1.30$, *ns*; Block 3 $\chi^2(1) = 1.44$, *ns*.

Table 35. Linear Regression Analysis: Overall School by Targeting Individual and Systemic Factors

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.45	$F(5, 155) = 25.45^{***}$
Age (years)	-0.16	0.04	-0.26	-4.29	0.00		
? Aboriginal	0.16	0.10	0.10	1.68	0.10		
Prince George	0.35	0.15	0.15	2.40	0.02		
SAVRY	-0.02	0.01	-0.21	-3.27	0.00		
School start	0.70	0.10	0.45	7.08	0.00		
Block 2						0.02	$F(1, 154) = 4.30^{*}$
Age (years)	-0.13	0.04	-0.22	-3.37	0.00		
? Aboriginal	0.15	0.09	0.10	1.63	0.10		
Prince George	0.34	0.14	0.14	2.37	0.02		
SAVRY	-0.02	0.01	-0.20	-3.20	0.00		
School start	0.70	0.10	0.46	7.25	0.00		
Best practice 4	0.11	0.05	0.13	2.07	0.04		
Block 3						0.00	$F(1, 153) = 0.36$
Age (years)	-0.14	0.04	-0.22	-3.40	0.00		
? Aboriginal	0.15	0.09	0.10	1.61	0.11		
Prince George	0.34	0.14	0.14	2.39	0.02		
SAVRY	-0.02	0.01	-0.20	-3.17	0.00		
School start	0.71	0.10	0.46	7.26	0.00		
Best practice 4	0.11	0.05	0.13	2.10	0.04		
Age x BP 4	0.02	0.04	0.04	0.60	0.55		

Note. Sample size $n = 161$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 36. Linear Regression Analysis: Overall School by Community Involvement

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.45	$F(5, 155) = 25.45^{***}$
Age (years)	-0.16	0.04	-0.26	-4.29	0.00		
? Aboriginal	0.16	0.10	0.10	1.68	0.10		
Prince George	0.35	0.15	0.15	2.40	0.02		
SAVRY	-0.02	0.01	-0.21	-3.27	0.00		
School start	0.70	0.10	0.45	7.08	0.00		
Block 2						0.01	$F(1, 154) = 4.02^{*}$
Age (years)	-0.14	0.04	-0.22	-3.55	0.00		
? Aboriginal	0.14	0.09	0.09	1.52	0.13		
Prince George	0.31	0.14	0.13	2.14	0.03		
SAVRY	-0.02	0.01	-0.21	-3.31	0.00		
School start	0.70	0.10	0.46	7.23	0.00		
Best practice 5	0.09	0.04	0.12	2.00	0.05		
Block 3						0.00	$F(1, 153) = 0.21$
Age (years)	-0.14	0.04	-0.23	-3.54	0.00		
? Aboriginal	0.14	0.10	0.09	1.48	0.14		
Prince George	0.30	0.14	0.13	2.10	0.04		
SAVRY	-0.02	0.01	-0.21	-3.32	0.00		
School start	0.71	0.10	0.46	7.23	0.00		
Best practice 5	0.09	0.04	0.12	1.98	0.05		
Age x BP 5	0.02	0.04	0.03	0.46	0.65		

Note. Sample size $n = 161$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 37. Linear Regression Analysis: Overall School by Family Contact

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.45	$F(5, 155) = 25.45^{***}$
Age (years)	-0.16	0.04	-0.26	-4.29	0.00		
? Aboriginal	0.16	0.10	0.10	1.68	0.10		
Prince George	0.35	0.15	0.15	2.40	0.02		
SAVRY	-0.02	0.01	-0.21	-3.27	0.00		
School start	0.70	0.10	0.45	7.08	0.00		
Block 2						0.02	$F(1, 154) = 5.42^{*}$
Age (years)	-0.12	0.04	-0.20	-3.07	0.00		
? Aboriginal	0.15	0.09	0.10	1.62	0.12		
Prince George	0.31	0.14	0.13	2.19	0.03		
SAVRY	-0.02	0.01	-0.19	-3.03	0.00		
School start	0.71	0.10	0.46	7.28	0.00		
Best practice 6	0.10	0.04	0.15	2.33	0.02		
Block 3						0.00	$F(1, 153) = 0.39$
Age (years)	-0.13	0.04	-0.21	-3.10	0.00		
? Aboriginal	0.15	0.09	0.10	1.58	0.12		
Prince George	0.32	0.14	0.13	2.21	0.03		
SAVRY	-0.02	0.01	-0.19	-3.01	0.00		
School start	0.71	0.10	0.46	7.30	0.00		
Best practice 6	0.11	0.04	0.15	2.36	0.02		
Age x BP 6	0.02	0.03	0.04	0.63	0.53		

Note. Sample size $n = 161$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 38. Linear Regression Analysis: Overall School by Prioritizing School/Employment Re-entry

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.45	$F(5, 155) = 25.45$***
Age (years)	-0.16	0.04	-0.26	-4.29	0.00		
? Aboriginal	0.16	0.10	0.10	1.68	0.10		
Prince George	0.35	0.15	0.15	2.40	0.02		
SAVRY	-0.02	0.01	-0.21	-3.27	0.00		
School start	0.70	0.10	0.45	7.08	0.00		
Block 2						0.02	$F(1, 154) = 4.20$*
Age (years)	-0.15	0.04	-0.25	-4.14	0.00		
? Aboriginal	0.15	0.09	0.10	1.57	0.12		
Prince George	0.34	0.14	0.14	2.33	0.02		
SAVRY	-0.02	0.01	-0.20	-3.27	0.00		
School start	0.71	0.10	0.46	7.28	0.00		
Best practice 7	0.08	0.04	0.12	2.05	0.04		

Note. Sample size $n = 161$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 39. Logistic Regression Analysis: School at the End of Follow up by Targeting Individual and Systemic Factors

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Age (years)	-0.52	0.16	10.48	1	0.00	0.60	0.44 – 0.82
MH issues	-0.22	0.50	0.20	1	0.66	0.80	0.30 – 2.12
Prince George	1.19	0.64	3.43	1	0.06	3.28	0.93 – 11.56
SAVRY	-0.08	0.03	9.78	1	0.00	0.92	0.88 – 0.97
School start	1.06	0.38	7.85	1	0.00	2.90	1.38 – 6.10
Block 2							
Age (years)	-0.49	0.17	8.52	1	0.00	0.61	0.44 – 0.85
MH issues	-0.24	0.50	0.22	1	0.64	0.79	0.30 – 2.10
Prince George	1.18	0.64	3.38	1	0.07	3.24	0.92 – 11.37
SAVRY	-0.08	0.03	9.43	1	0.00	0.92	0.88 – 0.97
School start	1.07	0.38	7.92	1	0.00	2.92	1.38 – 6.16
Best practice 4	0.10	0.21	0.20	1	0.65	1.10	0.73 – 1.66
Block 3							
Age (years)	-0.48	0.17	7.70	1	0.01	0.62	0.44 – 0.87
MH issues	-0.25	0.50	0.25	1	0.62	0.78	0.29 – 2.08
Prince George	1.17	0.64	3.30	1	0.07	3.21	0.91 – 11.32
SAVRY	-0.08	0.03	9.39	1	0.00	0.92	0.88 – 0.97
School start	1.07	0.38	7.81	1	0.00	2.90	1.38 – 6.13
Best practice 4	0.09	0.21	0.17	1	0.68	1.09	0.72 – 1.66
Age x BP 4	-0.07	0.17	0.18	1	0.68	0.93	0.68 – 1.29

Note. Sample size $n = 157$. Model -2 Log likelihood = 174.79, $\chi^2(7) = 42.34$, $p < 0.001$, Cox and Snell $R^2 = 0.24$, Nagelkerke $R^2 = 0.32$, Hosmer-Lemeshow $\chi^2(8) = 5.97$, $p = 0.65$; Block 2 $\chi^2(1) = 0.20$, *ns*; Block 3 $\chi^2(1) = 0.18$, *ns*.

Table 40. Logistic Regression Analysis: School at the End of Follow up by Community Involvement

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Age (years)	-0.52	0.16	10.48	1	0.00	0.60	0.44 – 0.82
MH issues	-0.22	0.50	0.20	1	0.66	0.80	0.30 – 2.12
Prince George	1.19	0.64	3.43	1	0.06	3.28	0.93 – 11.56
SAVRY	-0.08	0.03	9.78	1	0.00	0.92	0.88 – 0.97
School start	1.06	0.38	7.85	1	0.00	2.90	1.38 – 6.10
Block 2							
Age (years)	-0.51	0.17	9.14	1	0.00	0.60	0.43 – 0.84
MH issues	-0.22	0.50	0.20	1	0.65	0.80	0.30 – 2.12
Prince George	1.17	0.65	3.25	1	0.07	3.22	0.90 – 11.50
SAVRY	-0.08	0.03	9.67	1	0.00	0.92	0.88 – 0.97
School start	1.07	0.38	7.86	1	0.00	2.90	1.38 – 6.11
Best practice 5	0.03	0.18	0.03	1	0.86	1.03	0.73 – 1.46
Block 3							
Age (years)	-0.51	0.17	8.77	1	0.00	0.60	0.43 – 0.84
MH issues	-0.22	0.50	0.19	1	0.66	0.80	0.30 – 2.15
Prince George	1.17	0.65	3.24	1	0.08	3.22	0.90 – 11.49
SAVRY	-0.08	0.03	9.52	1	0.00	0.92	0.88 – 0.97
School start	1.07	0.38	7.86	1	0.00	2.90	1.38 – 6.12
Best practice 5	0.03	0.18	0.03	1	0.86	1.03	0.73 – 1.46
Age x BP 5	0.01	0.16	0.00	1	0.96	1.01	0.74 – 1.36

Note. Sample size n = 157. Model -2 Log likelihood = 175.14, $\chi^2(7) = 41.99$, $p < 0.001$, Cox and Snell R^2 = 0.24, Nagelkerke R^2 = 0.31, Hosmer-Lemeshow $\chi^2(8) = 7.40$, $p = 0.49$; Block 2 $\chi^2(1) = 0.03$, *ns*; Block 3 $\chi^2(1) = 0.00$, *ns*.

Table 41. Logistic Regression Analysis: School at the End of Follow up by Family Contact

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Age (years)	-0.52	0.16	10.48	1	0.00	0.60	0.44 – 0.82
MH issues	-0.22	0.50	0.20	1	0.66	0.80	0.30 – 2.12
Prince George	1.19	0.64	3.43	1	0.06	3.28	0.93 – 11.56
SAVRY	-0.08	0.03	9.78	1	0.00	0.92	0.88 – 0.97
School start	1.06	0.38	7.85	1	0.00	2.90	1.38 – 6.10
Block 2							
Age (years)	-0.44	0.17	6.50	1	0.01	0.64	0.46 – 0.90
MH issues	-0.30	0.51	0.34	1	0.56	0.74	0.28 – 2.00
Prince George	1.11	0.64	2.96	1	0.08	3.04	0.86 – 10.76
SAVRY	-0.08	0.03	8.49	1	0.00	0.93	0.88 – 0.98
School start	1.08	0.38	8.00	1	0.00	2.96	1.40 – 6.26
Best practice 6	0.21	0.19	1.26	1	0.26	1.23	0.86 – 1.78
Block 3							
Age (years)	-0.43	0.18	5.44	1	0.02	0.66	0.46 – 0.94
MH issues	-0.28	0.51	0.31	1	0.58	0.75	0.28 – 2.04
Prince George	1.12	0.65	2.98	1	0.08	3.06	0.86 – 10.90
SAVRY	-0.07	0.03	8.50	1	0.00	0.93	0.88 – 0.98
School start	1.08	0.38	7.86	1	0.00	2.94	1.38 – 6.23
Best practice 6	0.20	0.19	1.19	1	0.28	1.23	0.85 – 1.78
Age x BP 6	-0.09	0.14	0.38	1	0.54	0.92	0.69 – 1.21

Note. Sample size n = 157. Model -2 Log likelihood = 173.52, $\chi^2(7) = 43.61$, $p < 0.001$, Cox and Snell $R^2 = 0.24$, Nagelkerke $R^2 = 0.32$, Hosmer-Lemeshow $\chi^2(8) = 10.24$, $p = 0.25$; Block 2 $\chi^2(1) = 1.26$, *ns*; Block 3 $\chi^2(1) = 0.39$, *ns*.

Table 42. Logistic Regression Analysis: School at the End of Follow up by Prioritizing School/Employment Re-entry

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Age (years)	-0.52	0.16	10.48	1	0.00	0.60	0.44 – 0.82
MH issues	-0.22	0.50	0.20	1	0.66	0.80	0.30 – 2.12
Prince George	1.19	0.64	3.43	1	0.06	3.28	0.93 – 11.56
SAVRY	-0.08	0.03	9.78	1	0.00	0.92	0.88 – 0.97
School start	1.06	0.38	7.85	1	0.00	2.90	1.38 – 6.10
Block 2							
Age (years)	-0.52	0.16	10.46	1	0.00	0.59	0.43 – 0.81
MH issues	-0.22	0.50	0.20	1	0.66	0.80	0.30 – 2.12
Prince George	1.20	0.64	3.46	1	0.06	3.31	0.94 – 11.67
SAVRY	-0.08	0.03	9.79	1	0.00	0.92	0.88 – 0.97
School start	1.06	0.38	7.75	1	0.00	2.88	1.37 – 6.08
Best practice 7	-0.03	0.17	0.04	1	0.84	0.97	0.70 – 1.35

Note. Sample size $n = 157$. Model -2 Log likelihood = 175.14, $\chi^2(6) = 42.00$, $p < 0.001$, Cox and Snell $R^2 = 0.24$, Nagelkerke $R^2 = 0.31$, Hosmer-Lemeshow $\chi^2(8) = 6.21$, $p = 0.62$; Block 2 $\chi^2(1) = 0.04$, *ns*.

Table 43. Linear Regression Analysis: Number of Community Programs by Targeting Individual and Systemic Factors

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.19	$F(4, 155) = 8.95^{***}$
Addictions	0.34	0.24	0.11	1.42	0.16		
SAVRY	0.03	0.02	0.17	2.06	0.04		
File quality	0.55	0.21	0.21	2.64	0.01		
File time (months)	0.04	0.02	0.20	2.56	0.01		
Block 2						0.00	$F(1, 154) = 0.24$
Addictions	0.35	0.24	0.11	1.42	0.16		
SAVRY	0.03	0.02	0.17	2.06	0.04		
File quality	0.54	0.21	0.21	2.60	0.01		
File time (months)	0.04	0.02	0.19	2.27	0.02		
Best practice 4	0.06	0.13	0.04	0.49	0.63		
Block 3						0.00	$F(1, 153) = 0.68$
Addictions	0.36	0.24	0.12	1.47	0.14		
SAVRY	0.03	0.02	0.16	1.96	0.05		
File quality	0.57	0.21	0.22	2.70	0.01		
File time (months)	0.01	0.04	0.06	0.32	0.75		
Best practice 4	-0.14	0.28	-0.08	-0.49	0.62		
FT x BP 4	0.01	0.02	0.21	0.82	0.41		

Note. Sample size $n = 160$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 44. Linear Regression Analysis: Number of Community Programs by Community Involvement

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.19	$F(4, 155) = 8.95$***
Addictions	0.34	0.24	0.11	1.42	0.16		
SAVRY	0.03	0.02	0.17	2.06	0.04		
File quality	0.55	0.21	0.21	2.64	0.01		
File time (months)	0.04	0.02	0.20	2.56	0.01		
Block 2						0.01	$F(1, 154) = 2.21$
Addictions	0.34	0.24	0.11	1.41	0.16		
SAVRY	0.03	0.02	0.16	2.04	0.04		
File quality	0.52	0.21	0.20	2.53	0.01		
File time (months)	0.03	0.02	0.16	1.89	0.06		
Best practice 5	0.17	0.11	0.12	1.49	0.14		
Block 3						0.00	$F(2, 152) = 0.49$
Addictions	0.36	0.24	0.12	1.48	0.14		
SAVRY	0.03	0.02	0.16	2.03	0.04		
File quality	0.90	0.43	0.35	2.08	0.04		
File time (months)	0.02	0.04	0.07	0.38	0.71		
Best practice 5	0.23	0.26	0.16	0.89	0.38		
FQ x BP 5	0.01	0.02	0.14	0.50	0.62		
FT x BP 5	-0.17	0.17	-0.22	-0.98	0.33		

Note. Sample size $n = 160$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 45. Linear Regression Analysis: Number of Community Programs by Family Involvement

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.19	$F(4, 155) = 8.95$***
Addictions	0.34	0.24	0.11	1.42	0.16		
SAVRY	0.03	0.02	0.17	2.06	0.04		
File quality	0.55	0.21	0.21	2.64	0.01		
File time (months)	0.04	0.02	0.20	2.56	0.01		
Block 2						0.00	$F(1, 154) = 0.08$
Addictions	0.34	0.24	0.11	1.40	0.16		
SAVRY	0.03	0.02	0.17	2.06	0.04		
File quality	0.55	0.21	0.21	2.62	0.01		
File time (months)	0.04	0.02	0.20	2.37	0.02		
Best practice 6	0.03	0.11	0.02	0.28	0.78		
Block 3						0.00	$F(1, 153) = 0.90$
Addictions	0.37	0.24	0.12	1.50	0.14		
SAVRY	0.03	0.02	0.16	1.92	0.06		
File quality	0.57	0.21	0.22	2.72	0.01		
File time (months)	0.01	0.03	0.07	0.43	0.68		
Best practice 6	-0.18	0.25	-0.13	-0.73	0.47		
FT x BP 6	0.01	0.01	0.23	0.95	0.34		

Note. Sample size $n = 160$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 46. Linear Regression Analysis: Number of Community Programs by Prioritizing School/Employment Re-entry

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.19	$F(4, 155) = 8.95^{***}$
Addictions	0.34	0.24	0.11	1.42	0.16		
SAVRY	0.03	0.02	0.17	2.06	0.04		
File quality	0.55	0.21	0.21	2.64	0.01		
File time (months)	0.04	0.02	0.20	2.56	0.01		
Block 2						0.00	$F(1, 154) = 0.01$
Addictions	0.35	0.24	0.11	0.16	0.16		
SAVRY	0.03	0.02	0.17	0.04	0.04		
File quality	0.55	0.21	0.21	0.01	0.01		
File time (months)	0.04	0.02	0.21	0.02	0.02		
Best practice 7	-0.01	0.12	-0.01	0.94	0.94		
Block 3						0.02	$F(1, 151) = 0.95$
Addictions	0.70	0.48	0.22	1.46	0.15		
SAVRY	0.03	0.02	0.17	2.10	0.04		
File quality	1.00	0.39	0.39	2.59	0.01		
File time (months)	0.03	0.04	0.16	0.90	0.37		
Best practice 7	0.28	0.30	0.20	0.93	0.35		
Addictions x BP 7	-0.17	0.21	-0.14	-0.79	0.43		
FQ x BP 7	-0.25	0.19	-0.30	-1.35	0.18		
FT x BP 7	0.00	0.02	0.08	0.27	0.79		

Note. Sample size $n = 160$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 47. Logistic Regression Analysis: Any Recidivism by Targeting Individual and Systematic Factors

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.78	0.35	4.90	1	0.03	2.19	1.09 – 4.39
MH issues	0.48	0.47	1.06	1	0.30	1.62	0.65 – 4.04
Low cognitive	0.53	0.43	1.49	1	0.22	1.69	0.73 – 3.94
SAVRY	0.06	0.02	6.20	1	0.01	1.06	1.01 – 1.10
Block 2							
No Aboriginal	0.76	0.36	4.54	1	0.03	2.14	1.06 – 4.29
MH issues	0.44	0.47	0.89	1	0.35	1.56	0.62 – 3.92
Low cognitive	0.60	0.44	1.89	1	0.17	1.83	0.77 – 4.32
SAVRY	0.05	0.02	6.04	1	0.01	1.06	1.01 – 1.10
Best practice 4	-0.21	0.19	1.29	1	0.26	0.81	0.56 – 1.17

Note. Sample size n = 161. Model -2 Log likelihood = 199.99, $\chi^2(5) = 20.45$, $p < 0.01$, Cox and Snell R^2 = 0.12, Nagelkerke R^2 = 0.16, Hosmer-Lemeshow $\chi^2(8) = 11.21$, $p = 0.19$; Block 2 $\chi^2(1) = 1.31$, *ns*.

Table 48. Logistic Regression Analysis: Any Recidivism by Community Involvement

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.78	0.35	4.90	1	0.03	2.19	1.09 – 4.39
MH issues	0.48	0.47	1.06	1	0.30	1.62	0.65 – 4.04
Low cognitive	0.53	0.43	1.49	1	0.22	1.69	0.73 – 3.94
SAVRY	0.06	0.02	6.20	1	0.01	1.06	1.01 – 1.10
Block 2							
No Aboriginal	0.74	0.36	4.26	1	0.04	2.09	1.04 – 4.22
MH issues	0.42	0.47	0.79	1	0.38	1.52	0.60 – 3.83
Low cognitive	0.64	0.44	2.07	1	0.15	1.89	0.80 – 4.50
SAVRY	0.06	0.02	6.46	1	0.01	1.06	1.01 – 1.10
Best practice 5	-0.24	0.16	2.37	1	0.12	0.78	0.57 – 1.07

Note. Sample size n = 161. Model -2 Log likelihood = 198.88, $\chi^2(5) = 21.57$, $p < 0.01$, Cox and Snell R^2 = 0.12, Nagelkerke R^2 = 0.17, Hosmer-Lemeshow $\chi^2(8) = 11.04$, $p = 0.20$; Block 2 $\chi^2(1) = 2.42$, *ns*.

Table 49. Logistic Regression Analysis: Any Recidivism by Family Contact

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.78	0.35	4.90	1	0.03	2.19	1.09 – 4.39
MH issues	0.48	0.47	1.06	1	0.30	1.62	0.65 – 4.04
Low cognitive	0.53	0.43	1.49	1	0.22	1.69	0.73 – 3.94
SAVRY	0.06	0.02	6.20	1	0.01	1.06	1.01 – 1.10
Block 2							
No Aboriginal	0.77	0.36	4.68	1	0.03	2.16	1.08 – 4.33
MH issues	0.47	0.47	0.98	1	0.32	1.59	0.64 – 4.00
Low cognitive	0.58	0.44	1.75	1	0.19	1.79	0.76 – 4.24
SAVRY	0.05	0.02	5.81	1	0.02	1.06	1.01 – 1.10
Best practice 6	-0.11	0.16	0.46	1	0.50	0.90	0.66 – 1.22

Note. Sample size n = 161. Model -2 Log likelihood = 200.84, $\chi^2(5) = 19.61$, $p < 0.01$, Cox and Snell R^2 = 0.12, Nagelkerke R^2 = 0.15, Hosmer-Lemeshow $\chi^2(8) = 4.48$, $p = 0.81$; Block 2 $\chi^2(1) = 0.47$, *ns*.

Table 50. Logistic Regression Analysis: Any Recidivism by Prioritizing School/Employment Re-entry

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.78	0.35	4.90	1	0.03	2.19	1.09 – 4.39
MH issues	0.48	0.47	1.06	1	0.30	1.62	0.65 – 4.04
Low cognitive	0.53	0.43	1.49	1	0.22	1.69	0.73 – 3.94
SAVRY	0.06	0.02	6.20	1	0.01	1.06	1.01 – 1.10
Block 2							
No Aboriginal	0.80	0.36	5.06	1	0.02	2.23	1.11 – 4.48
MH issues	0.51	0.47	1.17	1	0.28	1.66	0.66 – 4.19
Low cognitive	0.50	0.43	1.32	1	0.25	1.65	0.70 – 3.85
SAVRY	0.06	0.02	6.20	1	0.01	1.06	1.01 – 1.10
Best practice 7	0.10	0.16	0.46	1	0.50	1.11	0.82 – 1.50

Note. Sample size n = 161. Model -2 Log likelihood = 200.84, $\chi^2(5) = 19.61$, $p < 0.01$, Cox and Snell R^2 = 0.12, Nagelkerke R^2 = 0.15, Hosmer-Lemeshow $\chi^2(8) = 11.95$, $p = 0.15$; Block 2 $\chi^2(1) = 0.46$, *ns*.

Table 51. Logistic Regression Analysis: Any Recidivism by Best Practice Intensity

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.78	0.35	4.90	1	0.03	2.19	1.09 – 4.39
MH issues	0.48	0.47	1.06	1	0.30	1.62	0.65 – 4.04
Low cognitive	0.53	0.43	1.49	1	0.22	1.69	0.73 – 3.94
SAVRY	0.06	0.02	6.20	1	0.01	1.06	1.01 – 1.10
Block 2							
No Aboriginal	0.75	0.36	4.40	1	0.04	2.12	1.05 – 4.28
MH issues	0.44	0.47	0.86	1	0.36	1.55	0.61 – 3.90
Low cognitive	0.68	0.45	2.33	1	0.13	1.98	0.82 – 4.75
SAVRY	0.06	0.02	6.71	1	0.01	1.06	1.01 – 1.11
Best practice 8	-0.30	0.18	2.72	1	0.10	0.74	0.52 – 1.06
Block 3							
No Aboriginal	0.76	0.36	4.35	1	0.04	2.13	1.05 – 4.33
MH issues	0.35	0.48	0.54	1	0.46	1.42	0.55 – 3.64
Low cognitive	1.04	0.51	4.23	1	0.04	2.84	1.05 – 7.69
SAVRY	0.06	0.02	7.17	1	0.01	1.06	1.02 – 1.11
Best practice 8	-0.30	0.18	2.58	1	0.11	0.74	0.52 – 1.07
LC x BP 8	-0.30	0.18	2.74	1	0.10	0.74	0.51 – 1.06

Note. Sample size $n = 161$. Model -2 Log likelihood = 195.52, $\chi^2(6) = 24.93$, $p < 0.001$, Cox and Snell $R^2 = 0.14$, Nagelkerke $R^2 = 0.19$, Hosmer-Lemeshow $\chi^2(8) = 14.30$, $p = 0.07$; Block 2 $\chi^2(1) = 2.83$, *ns*; Block 3 $\chi^2(1) = 2.96$, *ns.*

Table 52. Logistic Regression Analysis: Any Recidivism by Youth Justice ISSP Intensity

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.78	0.35	4.90	1	0.03	2.19	1.09 – 4.39
MH issues	0.48	0.47	1.06	1	0.30	1.62	0.65 – 4.04
Low cognitive	0.53	0.43	1.49	1	0.22	1.69	0.73 – 3.94
SAVRY	0.06	0.02	6.20	1	0.01	1.06	1.01 – 1.10
Block 2							
No Aboriginal	0.79	0.36	4.90	1	0.03	2.20	1.10 – 4.41
MH issues	0.45	0.47	0.92	1	0.34	1.57	0.63 – 3.94
Low cognitive	0.59	0.44	1.81	1	0.18	1.81	0.76 – 4.27
SAVRY	0.06	0.02	6.55	1	0.01	1.06	1.01 – 1.11
YJ ISSP 8	-0.16	0.17	0.82	1	0.37	0.86	0.61 – 1.20
Block 3							
No Aboriginal	0.80	0.36	4.84	1	0.03	2.20	1.09 – 4.46
MH issues	0.38	0.48	0.65	1	0.42	1.47	0.58 – 3.74
Low cognitive	0.94	0.50	3.51	1	0.06	2.56	0.96 – 6.82
SAVRY	0.06	0.02	5.97	1	0.01	1.06	1.01 – 1.11
YJ ISSP 8	-0.16	0.20	0.88	1	0.35	0.85	0.60 – 1.20
LC x YJ ISSP 8	-0.31	0.20	2.58	1	0.11	0.73	0.50 – 1.07

Note. Sample size $n = 161$. Model -2 Log likelihood = 197.65, $\chi^2(6) = 22.80$, $p < 0.01$, Cox and Snell $R^2 = 0.13$, Nagelkerke $R^2 = 0.18$, Hosmer-Lemeshow $\chi^2(8) = 5.80$, $p = 0.67$; Block 2 $\chi^2(1) = 0.83$, *ns*; Block 3 $\chi^2(1) = 2.83$, *ns.*

Table 53. Logistic Regression Analysis: Any Recidivism by Duration

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.78	0.35	4.90	1	0.03	2.19	1.09 – 4.39
MH issues	0.48	0.47	1.06	1	0.30	1.62	0.65 – 4.04
Low cognitive	0.53	0.43	1.49	1	0.22	1.69	0.73 – 3.94
SAVRY	0.06	0.02	6.20	1	0.01	1.06	1.01 – 1.10
Block 2							
No Aboriginal	0.77	0.36	4.69	1	0.03	2.16	1.08 – 4.34
MH issues	0.47	0.47	1.02	1	0.31	1.60	0.64 – 4.02
Low cognitive	0.57	0.44	1.72	1	0.19	1.77	0.75 – 4.18
SAVRY	0.06	0.02	6.26	1	0.01	1.06	1.01 – 1.10
Best Practice 9	-0.12	0.18	0.41	1	0.52	0.89	0.62 – 1.27

Note. Sample size n = 161. Model -2 Log likelihood = 200.89, $\chi^2(5) = 19.56$, $p < 0.01$, Cox and Snell R^2 = 0.11, Nagelkerke R^2 = 0.15, Hosmer-Lemeshow $\chi^2(8) = 5.46$, $p = 0.71$; Block 2 $\chi^2(1) = 0.41$, *ns*.

Table 54. Logistic Regression Analysis: Any Recidivism by Number of Hours

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.78	0.35	4.90	1	0.03	2.19	1.09 – 4.39
MH issues	0.48	0.47	1.06	1	0.30	1.62	0.65 – 4.04
Low cognitive	0.53	0.43	1.49	1	0.22	1.69	0.73 – 3.94
SAVRY	0.06	0.02	6.20	1	0.01	1.06	1.01 – 1.10
Block 2							
No Aboriginal	0.74	0.36	4.18	1	0.04	2.10	1.03 – 4.27
MH issues	0.37	0.48	0.60	1	0.44	1.45	0.56 – 3.73
Low cognitive	0.80	0.46	3.01	1	0.08	2.22	0.90 – 5.49
SAVRY	0.06	0.02	7.09	1	0.01	1.06	1.02 – 1.11
# of hours	-0.01	0.00	5.50	1	0.02	0.99	0.99 – 1.00
Block 3							
No Aboriginal	0.73	0.37	3.95	1	0.05	2.07	1.01 – 4.24
MH issues	0.32	0.49	0.42	1	0.52	1.37	0.53 – 3.57
Low cognitive	1.08	0.50	4.73	1	0.01	2.95	1.11 – 7.79
SAVRY	0.06	0.02	7.28	1	0.03	1.06	1.02 – 1.11
# of hours	-0.01	0.00	4.75	1	0.08	0.99	0.99 – 1.00
LC x # of hours	-0.01	0.01	3.02	1	0.00	0.99	0.98 – 1.00

Note. Sample size n = 161. Model -2 Log likelihood = 190.88, $\chi^2(6) = 29.57$, $p < 0.001$, Cox and Snell R^2 = 0.17, Nagelkerke R^2 = 0.22, Hosmer-Lemeshow $\chi^2(8) = 6.11$, $p = 0.63$; Block 2 $\chi^2(1) = 6.42$, $p < 0.05$; Block 3 $\chi^2(1) = 4.00$, $p < 0.05$.

Table 55. Logistic Regression Analysis: Any Recidivism by the Number of Hours for Youth with and without Low Cognitive Functioning

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Low cognitive functioning							
SAVRY total	0.10	0.06	2.63	1	0.10	1.10	0.98 – 1.24
# of hours	-0.02	0.01	7.61	1	0.01	0.98	0.96 – 0.99
No cognitive issues							
SAVRY total	0.06	0.02	6.26	1	0.01	1.06	1.01 – 1.11
# of hours	0.00	0.00	0.76	1	0.38	1.00	0.99 – 1.00

Note. Sample size for youth with low cognitive functioning $n = 34$. Model -2 Log likelihood = 30.92, $\chi^2(2) = 15.15$, $p < 0.001$, Cox and Snell $R^2 = 0.36$, Nagelkerke $R^2 = 0.48$, Hosmer-Lemeshow $\chi^2(8) = 10.37$, $p = 0.24$; Block 2 $\chi^2(1) = 13.21$, $p < 0.01$.

Sample size for youth with no identified cognitive functioning problems $n = 127$. Model -2 Log likelihood = 162.67, $\chi^2(2) = 7.60$, $p < 0.05$, Cox and Snell $R^2 = 0.06$, Nagelkerke $R^2 = 0.08$, Hosmer-Lemeshow $\chi^2(8) = 10.13$, $p = 0.26$; Block 2 $\chi^2(1) = 0.80$, *ns*.

Table 56. Logistic Regression Analysis: Any Recidivism by Number of Months

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.78	0.35	4.90	1	0.03	2.19	1.09 – 4.39
MH issues	0.48	0.47	1.06	1	0.30	1.62	0.65 – 4.04
Low cognitive	0.53	0.43	1.49	1	0.22	1.69	0.73 – 3.94
SAVRY	0.06	0.02	6.20	1	0.01	1.06	1.01 – 1.10
Block 2							
No Aboriginal	0.80	0.36	4.96	1	0.03	2.22	1.10 – 4.46
MH issues	0.48	0.47	1.05	1	0.31	1.62	0.64 – 4.07
Low cognitive	0.63	0.44	2.04	1	0.15	1.87	0.79 – 4.44
SAVRY	0.06	0.02	6.45	1	0.01	1.06	1.01 – 1.11
# of months	-0.03	0.02	1.77	1	0.18	0.97	0.93 – 1.02
Block 3							
No Aboriginal	0.81	0.36	5.08	1	0.02	2.26	1.11 – 4.57
MH issues	0.40	0.48	0.71	1	0.40	1.49	0.59 – 3.79
Low cognitive	0.95	0.50	3.63	1	0.06	2.59	0.97 – 6.90
SAVRY	0.06	0.02	7.05	1	0.01	1.06	1.02 – 1.11
# of months	-0.03	0.02	1.64	1	0.20	0.97	0.93 – 1.02
LC x # of months	-0.04	0.03	2.17	1	0.14	0.96	0.91 – 1.01

Note. Sample size n = 161. Model -2 Log likelihood = 197.20, $\chi^2(6) = 23.25$, $p < 0.01$, Cox and Snell R^2 = 0.13, Nagelkerke R^2 = 0.18, Hosmer-Lemeshow $\chi^2(8) = 9.65$, $p = 0.29$; Block 2 $\chi^2(1) = 1.81$, *ns*; Block 3 $\chi^2(1) = 2.30$, *ns.*

Table 57. Logistic Regression Analysis: Any Recidivism by the Number of Months in the Community

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.78	0.35	4.90	1	0.03	2.19	1.09 – 4.39
MH issues	0.48	0.47	1.06	1	0.30	1.62	0.65 – 4.04
Low cognitive	0.53	0.43	1.49	1	0.22	1.69	0.73 – 3.94
SAVRY	0.06	0.02	6.20	1	0.01	1.06	1.01 – 1.10
Block 2							
No Aboriginal	0.74	0.36	4.20	1	0.04	2.10	1.03 – 4.25
MH issues	0.48	0.48	1.03	1	0.31	1.62	0.64 – 4.12
Low cognitive	0.68	0.45	2.34	1	0.13	1.98	0.83 – 4.75
SAVRY	0.05	0.02	5.69	1	0.02	1.06	1.01 – 1.10
Community months	-0.06	0.02	4.91	1	0.03	0.95	0.90 – 0.99

Note. Sample size $n = 161$. Model -2 Log likelihood = 196.10, $\chi^2(5) = 24.35$, $p < 0.001$, Cox and Snell $R^2 = 0.14$, Nagelkerke $R^2 = 0.19$, Hosmer-Lemeshow $\chi^2(8) = 3.15$, $p = 0.92$; Block 2 $\chi^2(1) = 5.21$, $p < 0.05$.

Table 58. Frequency of Need Areas, Need Activities, and Appropriateness of Criminogenic Need Activities

	Rated C or D[a]	Activities by % youth	Appropriate Activities[b]	Over-targeted	Under-targeted
Need area	*n* (%)	*n* (%)	*n* (%)	*n* (%)	*n* (%)
Family relationships	143 (81.7)	20 (11.4)	48 (27.5)	2 (1.1)	125 (71.4)
Supervision	96 (54.9)	22 (12.6)	89 (50.9)	6 (3.4)	80 (45.7)
Housing	97 (55.4)	19 (10.9)	95 (54.3)	1 (0.6)	79 (45.1)
School/Employment	148 (84.6)	105 (60.0)	98 (56.0)	17 (9.7)	60 (34.3)
Peers	147 (84.0)	1 (0.6)	29 (16.6)	0 (0.0)	146 (83.4)
Substance use	137 (78.3)	24 (13.7)	58 (33.2)	2 (1.1)	115 (65.7)
Recreation	126 (72.0)	52 (29.7)	71 (40.6)	15 (8.5)	89 (50.9)
MH/Behavioural	158 (90.3)	42 (24.0)	53 (30.3)	3 (1.7)	119 (68.0)
Attitudes	127 (72.6)	31 (17.7)	75 (42.9)	2 (1.1)	98 (56.0)

Notes: Percentages reflect the total youth for which there was a community risk assessment on file (*n* = 175).

[a] Needs rated C were identified as moderately problematic for youth at the time of ISSP referral; needs rated D were identified as highly problematic .

[b] Activities were considered appropriate where a moderate or high need was targeted through program activities or where a strength or non-problematic area was not targeted.

Table 59. Logistic Regression Analysis: Any Recidivism by Criminogenic Need Activities

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.78	0.35	4.90	1	0.03	2.19	1.09 – 4.39
MH issues	0.48	0.47	1.06	1	0.30	1.62	0.65 – 4.04
Low cognitive	0.53	0.43	1.49	1	0.22	1.69	0.73 – 3.94
SAVRY	0.06	0.02	6.20	1	0.01	1.06	1.01 – 1.10
Block 2							
No Aboriginal	0.74	0.36	4.27	1	0.04	2.10	1.04 – 4.22
MH issues	0.43	0.47	0.85	1	0.36	1.54	0.61 – 3.88
Low cognitive	0.66	0.44	2.18	1	0.14	1.92	0.81 – 4.60
SAVRY	0.05	0.02	6.08	1	0.01	1.06	1.01 – 1.10
Best Practice 10	-0.28	0.17	2.65	1	0.10	0.76	0.54 – 1.06

Note. Sample size $n = 161$. Model -2 Log likelihood = 198.56, $\chi^2(5) = 21.88$, $p < 0.01$, Cox and Snell $R^2 = 0.13$, Nagelkerke $R^2 = 0.17$, Hosmer-Lemeshow $\chi^2(8) = 7.00$, $p = 0.54$; Block 2 $\chi^2(1) = 2.74$, *ns*.

Table 60. Logistic Regression Analysis: Any Recidivism by Specific Criminogenic Needs

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.78	0.35	4.90	1	0.03	2.19	1.09 – 4.39
MH issues	0.48	0.47	1.06	1	0.30	1.62	0.65 – 4.04
Low cognitive	0.53	0.43	1.49	1	0.22	1.69	0.73 – 3.94
SAVRY	0.06	0.02	6.20	1	0.01	1.06	1.01 – 1.10
Block 2							
No Aboriginal	0.78	0.36	4.84	1	0.03	2.19	1.09 – 4.40
MH issues	0.49	0.47	1.08	1	0.30	1.62	0.65 – 4.07
Low cognitive	0.58	0.44	1.78	1	0.18	1.79	0.76 – 4.20
SAVRY	0.06	0.02	6.49	1	0.01	1.06	1.01 – 1.10
Best Practice 11	-0.23	0.17	1.89	1	0.17	0.79	0.57 – 1.10

Note. Sample size n = 161. Model -2 Log likelihood = 199.38, $\chi^2(5) = 21.07$, $p < 0.01$, Cox and Snell R^2 = 0.12, Nagelkerke R^2 = 0.16, Hosmer-Lemeshow $\chi^2(8) = 7.29$, $p = 0.50$; Block 2 $\chi^2(1) = 1.93$, *ns.*

Table 61. Logistic Regression Analysis: Any Recidivism by the Balance between Support and Supervision Activities

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.78	0.35	4.90	1	0.03	2.19	1.09 – 4.39
MH issues	0.48	0.47	1.06	1	0.30	1.62	0.65 – 4.04
Low cognitive	0.53	0.43	1.49	1	0.22	1.69	0.73 – 3.94
SAVRY	0.06	0.02	6.20	1	0.01	1.06	1.01 – 1.10
Block 2							
No Aboriginal	0.77	0.36	4.74	1	0.03	2.17	1.08 – 4.36
MH issues	0.48	0.47	1.04	1	0.31	1.61	0.64 – 4.02
Low cognitive	0.54	0.43	1.55	1	0.21	1.72	0.73 – 4.01
SAVRY	0.06	0.02	6.03	1	0.01	1.06	1.01 – 1.10
Best Practice 12	-0.05	0.16	0.10	1	0.75	0.95	0.69 – 1.30

Note. Sample size n = 161. Model -2 Log likelihood = 201.20, $\chi^2(5) = 19.24$, $p < 0.01$, Cox and Snell R^2 = 0.11, Nagelkerke R^2 = 0.15, Hosmer-Lemeshow $\chi^2(8) = 15.01$, $p = 0.06$; Block 2 $\chi^2(1) = 0.10$, *ns.*

Table 62. Logistic Regression Analysis: Any Recidivism by Adherence to Youth Justice ISSP Guidelines

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp (*B*)
Block 1							
No Aboriginal	0.78	0.35	4.90	1	0.03	2.19	1.09 – 4.39
MH issues	0.48	0.47	1.06	1	0.30	1.62	0.65 – 4.04
Low cognitive	0.53	0.43	1.49	1	0.22	1.69	0.73 – 3.94
SAVRY	0.06	0.02	6.20	1	0.01	1.06	1.01 – 1.10
Block 2							
No Aboriginal	0.77	0.36	4.84	1	0.03	2.19	1.09 – 4.41
MH issues	0.48	0.47	1.06	1	0.30	1.62	0.65 – 4.04
Low cognitive	0.53	0.43	1.48	1	0.22	1.69	0.72 – 3.96
SAVRY	0.06	0.02	6.18	1	0.01	1.06	1.01 – 1.10
Best Practice 13	0.00	0.17	0.00	1	0.99	1.00	0.71 – 1.10

Note. Sample size n = 161. Model -2 Log likelihood = 201.30, $\chi^2(5) = 19.14$, $p < 0.01$, Cox and Snell R^2 = 0.11, Nagelkerke R^2 = 0.15, Hosmer-Lemeshow $\chi^2(8) = 17.93$, $p = 0.02$; Block 2 $\chi^2(1) = 0.00$, *ns.*

Table 63. Logistic Regression Analysis: Any Part-time Employment by Best Practice Fidelity

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Gender	-1.44	0.48	9.02	1	0.00	0.24	0.09 – 0.60
File time (months)	0.12	0.03	14.06	1	0.00	1.12	1.06 – 1.19
Block 2							
Gender	-1.49	0.49	9.38	1	0.00	0.22	0.09 – 0.58
File time (months)	0.09	0.04	5.96	1	0.02	1.09	1.02 – 1.17
Best practice index	0.15	0.09	2.76	1	0.10	1.16	0.97 – 1.39
Block 3							
Gender	-1.49	0.49	9.39	1	0.00	0.22	0.09 – 0.58
File time (months)	0.10	0.08	1.56	1	0.21	1.11	0.94 – 1.29
Best practice index	0.18	0.19	0.94	1	0.33	1.20	0.83 – 1.74
FT x BP index	0.00	0.02	0.04	1	0.84	1.00	0.98 – 1.02

Note. Sample size $n = 139$. Model -2 Log likelihood = 143.03, $\chi^2(4) = 27.30$, $p < 0.001$, Cox and Snell $R^2 = 0.18$, Nagelkerke $R^2 = 0.25$, Hosmer-Lemeshow $\chi^2(8) = 12.94$, $p = 0.11$; Block 2 $\chi^2(1) = 2.84$, *ns*; Block 3 $\chi^2(1) = 0.04$, *ns*.

Table 64. Logistic Regression Analyses: Any Full-time Employment by Best Practice Fidelity

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Gender	-1.29	0.70	3.45	1	0.06	0.28	0.07 – 1.07
Age (years)	0.59	0.20	8.39	1	0.00	1.81	1.21 – 2.70
FASD	-0.34	0.45	0.58	1	0.45	0.71	0.30 – 1.71
MH Issues	-0.90	0.62	2.11	1	0.15	0.41	0.12 – 1.37
SAVRY	-0.02	0.03	0.44	1	0.50	0.98	0.93 – 1.04
File time (months)	0.08	0.03	6.45	1	0.01	1.08	1.02 – 1.15
Block 2							
Gender	-1.35	0.70	3.74	1	0.05	0.26	0.07 – 1.02
Age (years)	0.54	0.21	6.71	1	0.01	1.72	1.14 – 2.60
FASD	-0.35	0.45	0.59	1	0.44	0.71	0.29 – 1.71
MH Issues	-0.93	0.62	2.24	1	0.14	0.39	0.12 – 1.34
SAVRY	-0.01	0.03	0.18	1	0.67	0.99	0.94 – 1.04
File time (months)	0.10	0.04	6.58	1	0.01	1.10	1.02 – 1.18
Best practice index	-0.09	0.10	0.76	1	0.39	0.92	0.75 – 1.12
Block 3							
Gender	-1.32	0.70	3.51	1	0.06	0.27	0.07 – 1.06
Age (years)	0.53	0.22	6.06	1	0.01	1.70	1.11 – 2.60
FASD	-0.37	0.47	0.62	1	0.43	0.69	0.27 – 1.74
MH Issues	-0.80	0.63	1.62	1	0.20	0.45	0.13 – 1.54
SAVRY	-0.01	0.03	0.26	1	0.61	0.99	0.93 – 1.04
File time (months)	0.02	0.08	0.08	1	0.78	1.02	0.87 – 1.20
Best practice index	-0.28	0.21	1.67	1	0.20	0.76	0.50 – 1.15
Age x BP index	0.09	0.08	1.34	1	0.25	1.09	0.94 – 1.27
FT x BP index	0.01	0.01	0.90	1	0.34	1.01	0.99 – 1.03

Note. Sample size $n = 128$. Model -2 Log likelihood = 131.68, $\chi^2(9) = 30.32$, $p < 0.001$, Cox and Snell $R^2 = 0.21$, Nagelkerke $R^2 = 0.29$, Hosmer-Lemeshow $\chi^2(8) = 9.30$, $p = 0.32$; Block 2 $\chi^2(1) = 0.76$, *ns*; Block 3 $\chi^2(2) = 1.74$, *ns*.

Table 65. Linear Regression Analysis: Overall Employment by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.36	$F(7, 121) = 9.54^{***}$
Gender	-0.27	0.12	-0.17	-2.18	0.03		
Age (years)	0.07	0.01	0.14	1.77	0.08		
Aboriginal	0.03	0.10	0.02	0.32	0.75		
Low cognitive	-0.07	0.12	-0.05	-0.61	0.54		
SAVRY	-0.01	0.01	-0.17	-2.18	0.03		
File time (months)	0.01	0.01	0.15	1.99	0.05		
Employment start	0.72	0.14	0.40	5.20	0.00		
Block 2						0.01	$F(1, 120) = 1.22$
Gender	-0.25	0.12	-0.16	-2.03	0.04		
Age (years)	0.09	0.04	0.17	2.06	0.04		
Aboriginal	0.03	0.10	0.02	0.30	0.77		
Low cognitive	-0.08	0.12	-0.05	-0.68	0.50		
SAVRY	-0.01	0.01	-0.18	-2.34	0.02		
File time (months)	0.01	0.01	0.09	0.97	0.34		
Employment start	0.71	0.14	0.40	5.13	0.00		
Best practice index	0.02	0.02	0.11	1.11	0.27		
Block 3						0.00	$F(2, 118) = 0.32$
Gender	-0.25	0.12	-0.16	-2.01	0.05		
Age (years)	0.09	0.04	0.17	1.94	0.05		
Aboriginal	0.03	0.10	0.02	0.26	0.80		
Low cognitive	-0.07	0.12	-0.04	-0.57	0.57		
SAVRY	-0.01	0.01	-0.18	-2.34	0.02		
File time (months)	0.02	0.02	0.22	1.08	0.28		
Employment start	0.72	0.14	0.40	5.12	0.00		
Best practice index	0.05	0.04	0.23	1.20	0.23		
Age x BP index	0.00	0.02	0.01	0.10	0.92		
FT x BP index	0.00	0.00	-0.23	-0.73	0.47		

Note. Sample size $n = 129$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 66. Logistic Regression Analysis: Employment at the End of Follow up by Best Practice Fidelity

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
? Aboriginal	1.18	0.52	5.18	1	0.02	3.27	1.18 – 9.07
MH issues	-0.91	1.00	0.83	1	0.36	0.40	0.06 – 2.85
Addictions	-0.85	0.56	2.24	1	0.14	0.43	0.14 – 1.30
SAVRY	-0.06	0.03	4.00	1	0.05	0.94	0.88 – 1.00
File time (months)	0.10	0.04	7.21	1	0.01	1.10	1.03 – 1.19
Employment start	1.74	0.68	6.61	1	0.01	5.69	1.51 – 21.42
Block 2							
? Aboriginal	1.21	0.52	5.35	1	0.02	3.36	1.20 – 9.36
MH issues	-0.82	1.01	0.66	1	0.42	0.44	0.06 – 3.20
Addictions	-0.86	0.57	2.33	1	0.13	0.42	0.14 – 1.28
SAVRY	-0.07	0.03	4.32	1	0.04	0.93	0.87 – 1.00
File time (months)	0.09	0.04	3.94	1	0.05	1.09	1.00 – 1.19
Employment start	1.81	0.69	6.98	1	0.01	6.13	1.60 – 23.51
Best practice index	0.07	0.11	0.41	1	0.52	1.07	0.86 – 1.34
Block 3							
? Aboriginal	1.20	0.52	5.24	1	0.02	3.32	1.19 – 9.30
MH issues	-0.85	1.02	0.69	1	0.41	0.43	0.06 – 3.17
Addictions	-0.88	0.57	2.40	1	0.12	0.41	0.14 – 1.26
SAVRY	-0.07	0.03	4.32	1	0.04	0.93	0.87 – 1.00
File time (months)	0.11	0.10	1.14	1	0.29	1.12	0.91 – 1.36
Employment start	1.82	0.69	6.99	1	0.01	6.14	1.60 – 23.59
Best practice index	0.13	0.25	0.26	1	0.61	1.14	0.70 – 1.85
FT x BP index	0.00	0.01	0.06	1	0.80	1.00	0.97 – 1.02

Note. Sample size n = 123. Model -2 Log likelihood = 99.34, $\chi^2(8) = 39.54$, $p < 0.001$, Cox and Snell R^2 = 0.28, Nagelkerke R^2 = 0.40, Hosmer-Lemeshow $\chi^2(8) = 9.72$, $p = 0.29$; Block 2 $\chi^2(1) = 0.41$, *ns*; Block 3 $\chi^2(1) = 0.06$, *ns*.

Table 67. Logistic Regression Analysis: Any School by Best Practice Fidelity

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Age (years)	-0.73	0.34	4.64	1	0.03	0.48	0.25 – 0.94
Aboriginal	0.44	0.76	0.34	1	0.56	1.55	0.35 – 6.82
Victoria	-0.84	0.66	1.60	1	0.21	0.43	0.12 – 1.59
SAVRY	-0.14	0.05	6.56	1	0.01	0.87	0.78 – 0.97
File time (months)	0.09	0.04	4.13	1	0.04	1.09	1.00 – 1.19
Block 2							
Age (years)	-0.72	0.35	4.34	1	0.04	0.48	0.25 – 0.96
Aboriginal	0.44	0.76	0.34	1	0.56	1.56	0.35 – 6.88
Victoria	-0.85	0.67	1.62	1	0.20	0.43	0.12 – 1.58
SAVRY	-0.14	0.05	6.58	1	0.01	0.87	0.78 – 0.97
File time (months)	0.08	0.05	2.57	1	0.11	1.09	0.98 – 1.20
Best practice index	0.02	0.15	0.03	1	0.86	1.02	0.77 – 1.37
Block 3							
Age (years)	-0.87	0.40	4.64	1	0.03	0.42	0.19 – 0.92
Aboriginal	0.33	0.76	0.19	1	0.66	1.39	0.32 – 6.13
Victoria	-0.71	0.68	1.10	1	0.29	0.49	0.13 – 1.86
SAVRY	-0.15	0.06	6.70	1	0.01	0.86	0.77 – 0.96
File time (months)	-0.04	0.13	0.08	1	0.78	0.96	0.75 – 1.24
Best practice index	-0.24	0.34	0.52	1	0.47	0.78	0.40 – 1.52
Age x BP index	-0.06	0.16	0.17	1	0.68	0.94	0.69 – 1.28
FT x BP index	0.02	0.02	0.99	1	0.32	1.02	0.98 – 1.07

Note. Sample size n = 132. Model -2 Log likelihood = 70.74, $\chi^2(8) = 22.73$, $p < 0.01$, Cox and Snell R^2 = 0.16, Nagelkerke R^2 = 0.31, Hosmer-Lemeshow $\chi^2(8) = 8.82$, $p = 0.36$; Block 2 $\chi^2(1) = 0.03$, *ns*; Block 3 $\chi^2(2) = 1.37$, *ns*.

Table 68. Linear Regression Analysis: Overall School by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.17	*F*(5, 124) = 5.23***
Age (years)	0.12	0.04	0.23	2.73	0.01		
? Aboriginal	0.20	0.11	0.16	1.84	0.07		
Prince George	0.18	0.20	0.08	0.90	0.37		
SAVRY	-0.02	0.01	-0.27	-2.94	0.00		
School start	-0.24	0.12	-0.18	-2.08	0.04		
Block 2						0.02	*F*(1, 123) = 3.28
Age (years)	0.15	0.05	0.28	3.22	0.00		
? Aboriginal	0.19	0.11	0.14	1.70	0.09		
Prince George	0.15	0.20	0.06	0.78	0.44		
SAVRY	-0.02	0.01	-0.26	-3.06	0.00		
School start	-0.21	0.12	-0.16	-1.79	0.08		
Best practice index	0.04	0.02	0.16	1.81	0.07		
Block 3						0.00	*F*(1, 122) = 0.00
Age (years)	0.15	0.05	0.28	3.10	0.00		
? Aboriginal	0.19	0.11	0.14	1.69	0.09		
Prince George	0.15	0.20	0.06	0.77	0.44		
SAVRY	-0.02	0.01	-0.26	-3.05	0.00		
School start	-0.21	0.12	-0.16	-1.78	0.08		
Best practice index	0.04	0.02	0.16	1.80	0.07		
Age x BP index	0.00	0.02	0.00	-0.02	0.98		

Note. Sample size *n* = 131.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 69. Logistic Regression Analysis: School at the End of Follow up by Best Practice Fidelity

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
Age (years)	-0.51	0.18	8.36	1	0.00	0.60	0.43 – 0.85
MH issues	0.13	0.56	0.05	1	0.82	1.14	0.38 – 3.37
Prince George	1.04	0.78	1.80	1	0.18	2.83	0.62 – 12.94
SAVRY	-0.09	0.03	10.33	1	0.00	0.91	0.86 – 0.96
School start	0.70	0.41	2.92	1	0.09	2.01	0.90 – 4.48
Block 2							
Age (years)	-0.48	0.19	6.50	1	0.01	0.62	0.43 – 0.90
MH issues	0.13	0.56	0.05	1	0.82	1.14	0.38 – 3.37
Prince George	1.01	0.78	1.68	1	0.20	2.74	0.60 – 12.56
SAVRY	-0.09	0.03	10.37	1	0.00	0.91	0.86 – 0.96
School start	0.72	0.41	3.07	1	0.08	2.06	0.92 – 4.64
Best practice index	0.04	0.08	0.24	1	0.62	1.04	0.89 – 1.21
Block 3							
Age (years)	-0.42	0.19	4.84	1	0.03	0.66	0.45 – 0.96
MH issues	0.07	0.56	0.02	1	0.90	1.08	0.36 – 3.24
Prince George	0.98	0.78	1.57	1	0.21	2.66	0.58 – 12.29
SAVRY	-0.09	0.03	9.99	1	0.00	0.91	0.86 – 0.97
School start	0.71	0.42	2.91	1	0.09	2.04	0.90 – 4.62
Best practice index	0.04	0.08	0.24	1	0.63	1.04	0.89 – 1.20
Age x BP index	-0.08	0.07	1.62	1	0.20	0.92	0.81 – 1.05

Note. Sample size $n = 128$. Model -2 Log likelihood = 146.90, $\chi^2(7) = 30.51$, $p < 0.001$, Cox and Snell $R^2 = 0.21$, Nagelkerke $R^2 = 0.28$, Hosmer-Lemeshow $\chi^2(8) = 14.34$, $p = 0.07$; Block 2 $\chi^2(1) = 0.24$, *ns*; Block 3 $\chi^2(1) = 1.64$, *ns*.

Table 70. Linear Regression Analysis: Number of Community Programs by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.22	$F(4, 126) = 8.86^{***}$
Addictions	0.26	0.26	0.09	1.02	0.31		
SAVRY	0.03	0.02	0.14	1.60	0.11		
File quality	0.62	0.21	0.25	2.90	0.00		
File time (months)	0.05	0.02	0.24	2.82	0.01		
Block 2						0.02	$F(1, 125) = 2.64$
Addictions	0.25	0.26	0.08	0.96	0.34		
SAVRY	0.02	0.02	0.12	1.37	0.17		
File quality	0.60	0.21	0.25	2.83	0.00		
File time (months)	0.03	0.02	0.15	1.48	0.14		
Best practice index	0.08	0.05	0.16	1.63	0.11		
Block 3						0.01	$F(1, 123) = 0.78$
Addictions	0.27	0.26	0.09	1.04	0.30		
SAVRY	0.02	0.02	0.11	1.27	0.21		
File quality	0.75	0.52	0.30	1.45	0.15		
File time (months)	-0.02	0.05	-0.10	-0.45	0.65		
Best practice index	-0.02	0.12	-0.03	-0.15	0.88		
FQ x BP index	-0.01	0.08	-0.05	-0.18	0.86		
FT x BP index	0.01	0.01	0.42	1.24	0.22		

Note. Sample size $n = 131$.
*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 71. Correlations between Probation File Outcomes and Potential Covariates

	Probation file outcomes						
Covariates	1	2	3	4	5	6	7
Gender	0.08	0.20*	0.26**	0.34**	0.17	0.08	0.06
Age	0.04	-0.18*	-0.23**	-0.15	-0.14	-0.19*	-0.08
Aboriginal	0.13	0.23**	0.20*	0.26**	0.09	0.11	0.15
No Aboriginal	0.06	0.00	0.00	0.04	0.14	0.09	-0.06
? Aboriginal	-0.17	-0.20*	-0.17	-0.26**	-0.21*	-0.18*	-0.08
FASD	0.01	0.27**	0.26**	0.28**	0.26**	0.23*	0.17
MH issues	-0.11	-0.05	-0.04	0.05	0.12	0.09	0.20*
Low cognitive	0.03	0.13	0.13	0.02	0.27**	0.20*	0.07
Addictions	0.20*	0.27**	0.24**	0.20*	0.25*	0.34**	0.39**
Burnaby	-0.04	-0.12	0.04	-0.06	0.07	-0.02	-0.02
Victoria	-0.04	0.08	-0.05	0.07	-0.03	0.01	0.09
Prince George	0.14	0.08	-0.01	0.00	-0.07	0.02	-0.10
SAVRY	0.18*	0.53**	0.49**	0.48**	0.36**	0.42**	0.37**
File quality	0.16	0.47**	0.50**	0.35**	0.36**	0.36**	0.31**
File time	0.13	0.18*	0.19*	0.08	0.20**	0.11	0.14

Note. Sample size $n = 128$.

1. Failure to report to probation appointments.
2. Breach of conditions detected.
3. Curfew violations.
4. Periods of unknown whereabouts.
5. Police contact.
6. Incidences of violence.
7. Incidences of substance use.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 72. Linear Regression Analysis: Failure to Report to Probation by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.05	$F(2, 128) = 3.43$*
Addictions	0.19	0.14	0.12	1.38	0.17		
SAVRY	0.02	0.01	0.15	1.68	0.10		
Block 2						0.02	$F(1, 127) = 2.93$
Addictions	0.16	0.14	0.10	1.12	0.26		
SAVRY	0.01	0.01	0.14	1.61	0.11		
Best practice index	0.04	0.02	0.15	1.71	0.09		

Note. Sample size $n = 131$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 73. Linear Regression Analysis: Breach of Conditions by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.42	$F(9, 120) = 9.60^{***}$
Gender	0.20	0.15	0.10	1.35	0.18		
Age (years)	-0.01	0.05	-0.02	-0.23	0.82		
Aboriginal	0.08	0.16	0.04	0.48	0.63		
? Aboriginal	-0.06	0.13	-0.04	-0.45	0.65		
FASD	0.07	0.11	0.06	0.67	0.50		
Addictions	0.06	0.12	0.04	0.51	0.61		
SAVRY	0.04	0.01	0.39	4.64	0.00		
File quality	0.36	0.10	0.28	3.47	0.00		
File time (months)	0.01	0.01	0.12	1.56	0.12		
Block 2						0.00	$F(1, 119) = 0.64$
Gender	0.22	0.15	0.11	1.47	0.14		
Age (years)	0.00	0.05	0.01	0.10	0.92		
Aboriginal	0.06	0.16	0.03	0.37	0.72		
? Aboriginal	-0.07	0.13	-0.04	-0.52	0.60		
FASD	0.08	0.11	0.06	0.73	0.47		
Addictions	0.06	0.12	0.04	0.48	0.64		
SAVRY	0.04	0.01	0.38	4.41	0.00		
File quality	0.36	0.10	0.28	3.50	0.00		
File time (months)	0.01	0.01	0.08	0.82	0.42		
Best practice index	0.02	0.03	0.08	0.80	0.42		

Note. Sample size $n = 130$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

See next page for Block 3.

Table 73 (cont.). Linear Regression Analysis: Breach of Conditions by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 3						0.01	$F(3, 116) = 0.90$
Gender	0.23	0.15	0.12	1.54	0.13		
Age (years)	0.00	0.05	-0.01	-0.07	0.94		
Aboriginal	0.03	0.17	0.02	0.20	0.84		
? Aboriginal	-0.06	0.13	-0.04	-0.47	0.64		
FASD	0.10	0.11	0.07	0.90	0.37		
Addictions	0.06	0.12	0.04	0.51	0.61		
SAVRY	0.04	0.01	0.36	4.21	0.00		
File quality	0.14	0.25	0.11	0.57	0.57		
File time (months)	-0.01	0.02	-0.10	-0.51	0.61		
Best practice index	-0.07	0.06	-0.25	-1.13	0.26		
Age x BP index	0.02	0.02	0.06	0.78	0.44		
FQ x BP index	0.04	0.04	0.28	1.06	0.29		
FT x BP index	0.00	0.00	0.33	1.04	0.30		

Note. Sample size $n = 130$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 74. Linear Regression Analysis: Curfew Violations by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.42	$F(8, 120) = 10.72^{***}$
Gender	0.37	0.16	0.17	2.32	0.02		
Age (years)	-0.03	0.05	-0.05	-0.66	0.51		
Aboriginal	0.05	0.16	0.03	0.33	0.74		
FASD	0.10	0.12	0.07	0.86	0.39		
Addictions	0.04	0.13	0.02	0.31	0.76		
SAVRY	0.03	0.01	0.33	3.88	0.00		
File quality	0.46	0.11	0.33	4.11	0.00		
File time (months)	0.01	0.01	0.10	1.34	0.18		
Block 2						0.00	$F(1, 119) = 0.06$
Gender	0.36	0.16	0.17	2.23	0.03		
Age (years)	-0.04	0.06	-0.06	-0.70	0.48		
Aboriginal	0.06	0.16	0.03	0.36	0.72		
FASD	0.10	0.12	0.07	0.83	0.41		
Addictions	0.04	0.13	0.02	0.32	0.75		
SAVRY	0.03	0.01	0.33	3.84	0.00		
File quality	0.45	0.11	0.33	4.07	0.00		
File time (months)	0.01	0.01	0.11	1.24	0.22		
Best practice index	-0.01	0.03	-0.02	-0.24	0.81		

Note. Sample size $n = 129$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

See next page for Block 3.

Table 74 (cont.). Linear Regression Analysis: Curfew Violations by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 3						0.00	$F(3, 116) = 0.18$
Gender	0.36	0.16	0.17	2.19	0.03		
Age (years)	-0.04	0.06	-0.06	-0.69	0.49		
Aboriginal	0.05	0.16	0.03	0.33	0.74		
FASD	0.10	0.12	0.07	0.80	0.42		
Addictions	0.04	0.13	0.02	0.28	0.78		
SAVRY	0.03	0.01	0.33	3.75	0.00		
File quality	0.29	0.28	0.21	1.03	0.30		
File time (months)	0.02	0.02	0.13	0.60	0.55		
Best practice index	-0.04	0.07	-0.12	-0.54	0.59		
Age x BP index	0.00	0.02	0.00	0.00	0.99		
FQ x BP index	0.03	0.04	0.18	0.68	0.50		
FT x BP index	0.00	0.00	-0.01	-0.04	0.97		

Note. Sample size $n = 129$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 75. Linear Regression Analysis: Periods of Unknown Whereabouts by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.36	*F*(7, 121) = 9.93***
Gender	0.55	0.16	0.26	3.46	0.00		
Aboriginal	0.09	0.18	0.05	0.52	0.61		
? Aboriginal	-0.16	0.15	-0.09	-1.06	0.29		
FASD	0.10	0.12	0.07	0.84	0.40		
Addictions	-0.02	0.13	-0.01	-0.12	0.90		
SAVRY	0.03	0.01	0.32	3.69	0.00		
File quality	0.31	0.11	0.22	2.88	0.00		
Block 2						0.00	*F*(1, 120) = 0.10
Gender	0.55	0.16	0.26	3.45	0.00		
Aboriginal	0.09	0.18	0.04	0.49	0.63		
? Aboriginal	-0.16	0.15	-0.09	-1.07	0.29		
FASD	0.10	0.12	0.07	0.82	0.41		
Addictions	-0.02	0.13	-0.01	-0.16	0.88		
SAVRY	0.03	0.01	0.32	3.68	0.00		
File quality	0.30	0.11	0.22	2.75	0.01		
Best practice index	0.01	0.02	0.02	0.31	0.76		
Block 3						0.00	*F*(1, 119) = 0.10
Gender	0.55	0.16	0.26	3.43	0.00		
Aboriginal	0.09	0.18	0.04	0.49	0.62		
? Aboriginal	-0.16	0.15	-0.10	-1.08	0.28		
FASD	0.10	0.12	0.07	0.82	0.42		
Addictions	-0.02	0.13	-0.01	-0.14	0.88		
SAVRY	0.03	0.01	0.32	3.68	0.00		
File quality	0.38	0.25	0.27	1.49	0.14		
Best practice index	0.02	0.05	0.07	0.43	0.67		
FQ x BP index	-0.01	0.04	-0.08	-0.32	0.75		

Note. Sample size *n* = 129.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 76. Linear Regression Analysis: Police Contact by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.28	*F*(7, 122) = 6.77***
? Aboriginal	-0.14	0.11	-0.10	-1.20	0.23		
FASD	0.09	0.10	0.07	0.83	0.41		
Low cognitive	0.19	0.15	0.11	1.30	0.20		
Addictions	0.13	0.12	0.09	1.12	0.26		
SAVRY	0.02	0.01	0.19	2.08	0.04		
File quality	0.26	0.10	0.23	2.70	0.01		
File time (months)	0.01	0.01	0.14	1.68	0.10		
Block 2						0.01	*F*(1, 121) = 1.01
? Aboriginal	-0.14	0.11	-0.10	-1.21	0.23		
FASD	0.09	0.10	0.07	0.82	0.42		
Low cognitive	0.19	0.15	0.11	1.27	0.21		
Addictions	0.13	0.12	0.09	1.10	0.27		
SAVRY	0.02	0.01	0.18	1.92	0.06		
File quality	0.26	0.10	0.23	2.66	0.01		
File time (months)	0.01	0.01	0.09	0.86	0.39		
Best practice index	0.02	0.02	0.10	1.01	0.32		
Block 3						0.02	*F*(2, 119) = 1.74
? Aboriginal	-0.15	0.11	-0.11	-1.33	0.19		
FASD	0.06	0.10	0.05	0.60	0.55		
Low cognitive	0.23	0.15	0.14	1.54	0.13		
Addictions	0.10	0.12	0.08	0.90	0.37		
SAVRY	0.02	0.01	0.19	2.01	0.05		
File quality	0.03	0.23	0.02	0.11	0.91		
File time (months)	0.04	0.02	0.46	2.02	0.04		
Best practice index	0.07	0.06	0.27	1.20	0.23		
FQ x BP index	0.03	0.04	0.26	0.94	0.35		
FT x BP index	0.00	0.00	-0.61	-1.81	0.07		

Note. Sample size n = 130.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 77. Linear Regression Analysis: Violence by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.26	$F(7, 122) = 6.28^{***}$
Age (years)	-0.02	0.04	-0.04	-0.54	0.59		
? Aboriginal	-0.12	0.12	-0.09	-1.06	0.29		
FASD	0.09	0.11	0.08	0.84	0.40		
Low cognitive	0.21	0.15	0.12	1.42	0.16		
Addictions	0.17	0.12	0.12	1.42	0.16		
SAVRY	0.01	0.01	0.15	1.70	0.09		
File quality	0.31	0.09	0.27	3.25	0.00		
Block 2						0.02	$F(1, 121) = 2.79$
Age (years)	0.00	0.05	0.00	0.06	0.96		
? Aboriginal	-0.14	0.12	-0.10	-1.18	0.24		
FASD	0.09	0.11	0.08	0.85	0.40		
Low cognitive	0.19	0.15	0.11	1.28	0.20		
Addictions	0.14	0.12	0.10	1.21	0.23		
SAVRY	0.01	0.01	0.16	1.74	0.08		
File quality	0.28	0.10	0.25	2.95	0.00		
Best practice index	0.04	0.02	0.14	1.67	0.10		
Block 3						0.00	$F(2, 119) = 0.04$
Age (years)	0.01	0.05	0.01	0.15	0.88		
? Aboriginal	-0.14	0.12	-0.10	-1.16	0.25		
FASD	0.09	0.11	0.08	0.83	0.41		
Low cognitive	0.18	0.15	0.11	1.22	0.22		
Addictions	0.14	0.12	0.10	1.21	0.23		
SAVRY	0.01	0.01	0.15	1.69	0.09		
File quality	0.24	0.23	0.21	1.03	0.30		
Best practice index	0.03	0.04	0.11	0.61	0.54		
Age x BP index	-0.01	0.02	-0.03	-0.37	0.71		
FQ x BP index	0.01	0.04	0.06	0.21	0.83		

Note. Sample size $n = 130$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 78. Linear Regression Analysis: Substance Use by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.24	$F(4, 126) = 10.08$***
MH issues	0.18	0.17	0.09	1.10	0.27		
Addictions	0.42	0.13	0.27	3.22	0.00		
SAVRY	0.02	0.01	0.20	2.34	0.02		
File quality	0.23	0.10	0.18	2.21	0.03		
Block 2						0.02	$F(1, 125) = 4.19$*
MH issues	0.22	0.17	0.11	1.33	0.18		
Addictions	0.38	0.13	0.24	2.93	0.00		
SAVRY	0.02	0.01	0.20	2.34	0.02		
File quality	0.18	0.10	0.14	1.75	0.08		
Best practice index	0.05	0.02	0.16	2.05	0.04		
Block 3						0.01	$F(1, 124) = 1.24$
MH issues	0.23	0.17	0.11	1.36	0.18		
Addictions	0.39	0.13	0.24	2.97	0.00		
SAVRY	0.02	0.01	0.21	2.42	0.02		
File quality	0.42	0.24	0.34	1.77	0.08		
Best practice index	0.09	0.05	0.32	1.97	0.05		
FQ x BP index	-0.04	0.04	-0.29	-1.12	0.27		

Note. Sample size $n = 131$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 79. Logistic Regression Analysis: Any Recidivism by Best Practice Fidelity

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.51	0.40	1.58	1	0.21	1.66	0.75 – 3.65
MH issues	0.50	0.52	0.93	1	0.34	1.65	0.59 – 4.60
Low cognitive	0.49	0.46	1.14	1	0.29	1.64	0.66 – 4.03
SAVRY	0.07	0.03	6.54	1	0.01	1.07	1.02 – 1.12
Block 2							
No Aboriginal	0.45	0.41	1.24	1	0.26	1.57	0.71 – 3.49
MH issues	0.40	0.53	0.57	1	0.45	1.49	0.53 – 4.19
Low cognitive	0.62	0.48	1.69	1	0.19	1.86	0.73 – 4.76
SAVRY	0.07	0.03	7.25	1	0.01	1.07	1.02 – 1.13
Best practice index	-0.09	0.07	1.71	1	0.19	0.91	0.79 – 1.05

Note. Sample size $n = 132$. Model -2 Log likelihood = 162.57, $\chi^2(5) = 16.73$, $p < 0.01$, Cox and Snell $R^2 = 0.12$, Nagelkerke $R^2 = 0.16$, Hosmer-Lemeshow $\chi^2(8) = 9.98$, $p = 0.27$; Block 2 $\chi^2(1) = 1.74$, *ns*.

Table 80. Logistic Regression Analysis: Any Recidivism by Program-only Best Practice Fidelity

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.78	0.35	4.90	1	0.03	2.19	1.09 – 4.39
MH issues	0.48	0.47	1.06	1	0.30	1.62	0.65 – 4.04
Low cognitive	0.53	0.43	1.49	1	0.22	1.69	0.73 – 3.94
SAVRY	0.06	0.02	6.20	1	0.01	1.06	1.01 – 1.10
Block 2							
No Aboriginal	0.77	0.36	4.72	1	0.03	2.17	1.08 – 4.35
MH issues	0.45	0.47	0.92	1	0.34	1.57	0.63 – 3.93
Low cognitive	0.60	0.44	1.89	1	0.17	1.83	0.77 – 4.32
SAVRY	0.06	0.02	6.64	1	0.01	1.06	1.01 – 1.11
Program-only BP	-0.07	0.06	1.17	1	0.28	0.94	0.83 – 1.06

Note. Sample size n = 161. Model -2 Log likelihood = 200.12, $\chi^2(5) = 20.33$, $p < 0.01$, Cox and Snell R^2 = 0.12, Nagelkerke R^2 = 0.16, Hosmer-Lemeshow $\chi^2(8) = 12.33$, $p = 0.14$; Block 2 $\chi^2(1) = 1.18$, *ns*.

Table 81. Linear Regression Analysis: Convictions by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*B*	*t*	*p*	Δ R^2	*F* change
Block 1						0.19	$F(3, 124) = 9.64$***
Prince George	-1.44	0.66	-0.18	-2.20	0.03		
SAVRY	0.05	0.02	0.17	1.92	0.06		
Convictions pre-ISSP	0.23	0.07	0.30	3.36	0.00		
Block 2						0.01	$F(1, 123) = 1.03$
Prince George	-1.38	0.52	0.14	-2.10	0.04		
SAVRY	0.05	0.03	0.19	2.04	0.04		
Convictions pre-ISSP	0.22	0.07	0.30	3.24	0.00		
Best practice index	-0.07	0.07	-0.08	-1.01	0.31		

Note. Sample size $n = 128$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 82. Linear Regression Analysis: Offence Dates by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.21	$F(4, 122) = 8.20^{***}$
No Aboriginal	0.70	0.34	0.17	2.09	0.04		
FASD	0.32	0.26	0.10	1.20	0.23		
SAVRY	0.02	0.02	0.09	0.98	0.33		
Dates pre-ISSP	0.21	0.06	0.32	3.60	0.00		
Block 2						0.01	$F(1, 121) = 1.04$
No Aboriginal	0.67	0.34	0.16	1.98	0.05		
FASD	0.34	0.26	0.11	1.27	0.21		
SAVRY	0.02	0.02	0.10	1.08	0.28		
Dates pre-ISSP	0.20	0.06	0.31	3.49	0.00		
Best practice index	-0.06	0.05	-0.08	-1.02	0.31		

Note. Sample size $n = 127$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 83. Linear Regression Analysis: Non-breach Convictions by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*B*	*t*	*p*	Δ R^2	*F* change
Block 1						0.09	$F(3, 124) = 4.02$**
No Aboriginal	0.51	0.23	0.19	2.22	0.03		
SAVRY	0.02	0.01	0.15	1.72	0.09		
Non-breach pre-ISSP	0.07	0.07	0.10	1.04	0.30		
Block 2						0.02	$F(1, 123) = 3.19$
No Aboriginal	0.46	0.23	0.17	1.97	0.05		
SAVRY	0.02	0.01	0.16	1.84	0.07		
Non-breach pre-ISSP	0.09	0.07	0.12	1.33	0.19		
Best practice index	-0.07	0.04	-0.16	-1.79	0.08		

Note. Sample size $n = 128$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 84. Linear Regression Analysis: Violent Convictions by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.08	$F(3, 124) = 3.65$*
No Aboriginal	0.24	0.15	0.14	1.64	0.10		
SAVRY	0.01	0.01	0.08	0.86	0.39		
Violent pre-ISSP	0.11	0.05	0.20	2.18	0.03		
Block 2						0.01	$F(1, 123) = 1.55$
No Aboriginal	0.22	0.15	0.13	1.47	0.14		
SAVRY	0.01	0.01	0.09	0.96	0.34		
Violent pre-ISSP	0.12	0.05	0.21	2.31	0.02		
Best practice index	-0.03	0.02	-0.11	-1.24	0.22		

Note. Sample size $n = 128$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 85. Linear Regression Analysis: Indictable Convictions by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.09	*F*(3, 123) = 3.89*
Age (years)	0.16	0.07	0.20	2.30	0.02		
SAVRY	0.02	0.01	0.14	1.64	0.10		
Indictable pre-ISSP	0.18	0.09	0.18	2.08	0.04		
Block 2						0.02	*F*(1, 122) = 2.28
Age (years)	0.12	0.07	0.15	1.61	0.11		
SAVRY	0.02	0.01	0.15	1.74	0.08		
Indictable pre-ISSP	0.20	0.09	0.20	2.32	0.02		
Best practice index	-0.05	0.03	-0.14	-1.51	0.13		
Block 3						0.01	*F*(1, 121) = 0.76
Age (years)	0.14	0.08	0.17	1.79	0.08		
SAVRY	0.02	0.01	0.15	1.75	0.08		
Indictable pre-ISSP	0.20	0.09	0.20	2.30	0.02		
Best practice index	-0.05	0.03	-0.14	-1.53	0.13		
Age x BP index	-0.02	0.03	-0.08	-0.87	0.39		

Note. Sample size $n = 127$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 86. Linear Regression Analysis: Time in Community by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.09	$F(1, 130) = 13.38$***
SAVRY	-3.18	0.87	-0.30	-3.66	0.00		
Block 2						0.02	$F(1, 129) = 3.32$
SAVRY	-3.36	0.87	-0.32	-3.87	0.00		
Best practice index	4.61	2.53	0.15	1.82	0.07		

Note. Sample size $n = 132$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Time in community in the year prior to ISSP not available.

Table 87. Linear Regression Analysis: Seriousness Composite by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.15	$F(3, 124) = 7.08^{***}$
No Aboriginal	2.84	1.32	0.18	2.16	0.03		
SAVRY	0.12	0.08	0.13	1.51	0.13		
Seriousness pre-ISSP	0.17	0.06	0.23	2.62	0.01		
Block 2						0.03	$F(1, 123) = 4.25^{*}$
No Aboriginal	2.49	1.31	0.16	1.91	0.06		
SAVRY	0.13	0.08	0.15	1.68	0.10		
Seriousness pre-ISSP	0.18	0.06	0.25	2.84	0.00		
Best practice index	-0.44	0.21	-0.17	-2.06	0.04		

Note. Sample size $n = 128$.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 88. Logistic Regression Analysis: Any Recidivism Post-Follow up by Best Practice Fidelity

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Block 1							
No Aboriginal	0.84	0.46	3.40	1	0.06	2.31	0.95 – 5.65
SAVRY	0.07	0.03	6.11	1	0.01	1.07	1.01 – 1.13
Follow up time (months)	0.09	0.02	16.31	1	0.00	1.09	1.04 – 1.14
Block 2							
No Aboriginal	0.83	0.46	3.26	1	0.07	2.29	0.93 – 5.62
SAVRY total	0.07	0.03	6.17	1	0.01	1.07	1.02 – 1.13
Follow up time (months)	0.09	0.02	16.36	1	0.00	1.09	1.04 – 1.14
Best practice index	-0.02	0.07	0.05	1	0.82	0.98	0.85 – 1.13

Note. Sample size $n = 129$. Model -2 Log likelihood = 140.71, $\chi^2(4) = 32.43$, $p < 0.001$, Cox and Snell $R^2 = 0.22$, Nagelkerke $R^2 = 0.30$, Hosmer-Lemeshow $\chi^2(8) = 6.06$, $p = 0.64$; Block 2 $\chi^2(1) = 0.05$, *ns*.

Table 89. Linear Regression Analysis: Current Justice System Involvement by Best Practice Fidelity

Predictor	*B*	SE(*B*)	*β*	*t*	*p*	Δ R^2	*F* change
Block 1						0.08	*F*(3, 128) = 3.85*
No Aboriginal	0.27	0.16	0.18	1.69	0.09		
? Aboriginal	-0.12	0.15	-0.09	-0.79	0.43		
SAVRY	0.01	0.01	0.13	1.52	0.13		
Block 2						0.00	*F*(1, 127) = 0.01
No Aboriginal	0.27	0.16	0.18	1.66	0.10		
? Aboriginal	-0.12	0.15	-0.09	-0.80	0.43		
SAVRY	0.01	0.01	0.13	1.52	0.13		
Best practice index	0.00	0.02	-0.01	-0.10	0.92		

Note. Sample size *n* = 132.

*significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$.

Table 90. Logistic Regression Analysis: Any Recidivism by Best Practice Composite for Males and Females

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Males							
SAVRY	0.09	0.03	9.70	1	0.00	1.09	1.03 – 1.16
Best practice index	-0.06	0.07	0.69	1	0.41	0.94	0.81 – 1.09
Females							
SAVRY	0.14	0.09	2.12	1	0.15	1.15	0.95 – 1.38
Best practice index	-0.29	0.19	2.28	1	0.13	0.75	0.52 – 1.09

Note. Sample size for males $n = 104$. Model -2 Log likelihood = 131.84, $\chi^2(2) = 11.38$, $p < 0.01$, Cox and Snell $R^2 = 0.10$, Nagelkerke $R^2 = 0.14$, Hosmer-Lemeshow $\chi^2(8) = 10.00$, $p = 0.26$.

Sample size for females $n = 28$. Model -2 Log likelihood = 27.67, $\chi^2(2) = 5.83$, $p = 0.05$, Cox and Snell $R^2 = 0.19$, Nagelkerke $R^2 = 0.27$, Hosmer-Lemeshow $\chi^2(8) = 3.81$, $p = 0.80$.

Table 91. Logistic Regression Analysis: Any Recidivism by Best Practice Fidelity for Younger and Older Participants

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Younger participants							
SAVRY	0.17	0.06	9.67	1	0.00	1.18	1.06 – 1.32
Best practice index	-0.24	0.12	4.45	1	0.04	0.78	0.61 – 0.98
Older participants							
SAVRY	0.04	0.03	1.44	1	0.23	1.04	0.98 – 1.10
Best practice index	0.12	0.10	1.48	1	0.22	1.13	0.93 – 1.36

Note. Sample size for younger participants $n = 66$. Model -2 Log likelihood = 65.71, $\chi^2(2) = 20.81$, $p < 0.001$, Cox and Snell $R^2 = 0.27$, Nagelkerke $R^2 = 0.37$, Hosmer-Lemeshow $\chi^2(8) = 6.27$, $p = 0.51$.

Sample size for older participants $n = 66$. Model -2 Log likelihood = 87.77, $\chi^2(2) = 3.48$, *ns*, Cox and Snell $R^2 = 0.05$, Nagelkerke $R^2 = 0.07$, Hosmer-Lemeshow $\chi^2(8) = 15.05$, $p = 0.04$.

Table 92. Logistic Regression Analysis: Any Recidivism by Best Practice Fidelity for Aboriginal Identification

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Aboriginal							
SAVRY	0.21	0.12	2.97	1	0.08	1.24	0.97 – 1.57
Best practice index	-0.46	0.19	5.69	1	0.02	0.63	0.43 – 0.92
No Aboriginal							
SAVRY	0.06	0.05	1.90	1	0.17	1.06	0.97 – 1.16
Best practice index	0.01	0.12	0.00	1	0.96	1.01	0.80 – 1.27
? Aboriginal							
SAVRY	0.05	0.04	2.05	1	0.15	1.05	0.98 – 1.13
Best practice index	0.05	0.10	0.20	1	0.66	1.05	0.86 – 1.28

Note. Sample size for participants identified as Aboriginal $n = 32$. Model -2 Log likelihood = 26.92, $\chi^2(2) = 16.31$, $p < 0.001$, Cox and Snell $R^2 = 0.40$, Nagelkerke $R^2 = 0.54$, Hosmer-Lemeshow $\chi^2(8) = 6.25$, $p = 0.62$. Sample size for participants not identified as Aboriginal $n = 40$. Model -2 Log likelihood = 53.38, $\chi^2(2) = 2.08$, *ns*, Cox and Snell $R^2 = 0.05$, Nagelkerke $R^2 = 0.07$, Hosmer-Lemeshow $\chi^2(8) = 13.09$, $p = 0.11$. Sample size for participants with unknown Aboriginal identification $n = 60$. Model -2 Log likelihood = 75.91, $\chi^2(2) = 2.95$, *ns*, Cox and Snell $R^2 = 0.05$, Nagelkerke $R^2 = 0.07$, Hosmer-Lemeshow $\chi^2(8) = 6.10$, $p = 0.64$.

Table 93. Logistic Regression Analysis: Any Recidivism by Best Practice Fidelity for Higher- and Lower-risk Youth

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
Lower-risk youth							
SAVRY	0.12	0.05	5.04	1	0.02	1.12	1.02 – 1.24
Best practice index	0.02	0.09	0.06	1	0.80	1.02	0.86 – 1.22
Higher-risk youth							
SAVRY	0.12	0.08	2.21	1	0.14	1.12	0.96 – 1.31
Best practice index	-0.25	0.11	5.55	1	0.02	0.78	0.63 – 0.96

Note. Sample size for lower SAVRY risk $n = 68$. Model -2 Log likelihood = 80.39, $\chi^2(2) = 3.44$, $p < 0.05$, Cox and Snell $R^2 = 0.09$, Nagelkerke $R^2 = 0.13$, Hosmer-Lemeshow $\chi^2(8) = 6.66$, $p = 0.57$.

Sample size for higher SAVRY risk $n = 64$. Model -2 Log likelihood = 80.42, $\chi^2(2) = 8.31$, $p < 0.05$, Cox and Snell $R^2 = 0.12$, Nagelkerke $R^2 = 0.16$, Hosmer-Lemeshow $\chi^2(8) = 9.40$, $p = 0.31$.

Table 94. Logistic Regression Analysis: Any Recidivism by Best Practice Fidelity for Youth with and without Addictions

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp(*B*)
No Addictions							
SAVRY	0.11	0.04	8.56	1	0.00	1.12	1.04 - 1.21
Best practice index	-0.05	0.09	0.33	1	0.56	0.95	0.79 – 1.13
Addictions							
SAVRY	0.04	0.04	1.11	1	0.29	1.04	0.96 – 1.13
Best practice index	-0.16	0.11	2.18	1	0.14	0.86	0.69 – 1.05

Note. Sample size for no addictions $n = 68$. Model -2 Log likelihood = 78.63, $\chi^2(2) = 10.81$, $p < 0.01$, Cox and Snell $R^2 = 0.15$, Nagelkerke $R^2 = 0.20$, Hosmer-Lemeshow $\chi^2(8) = 11.46$, $p = 0.12$.

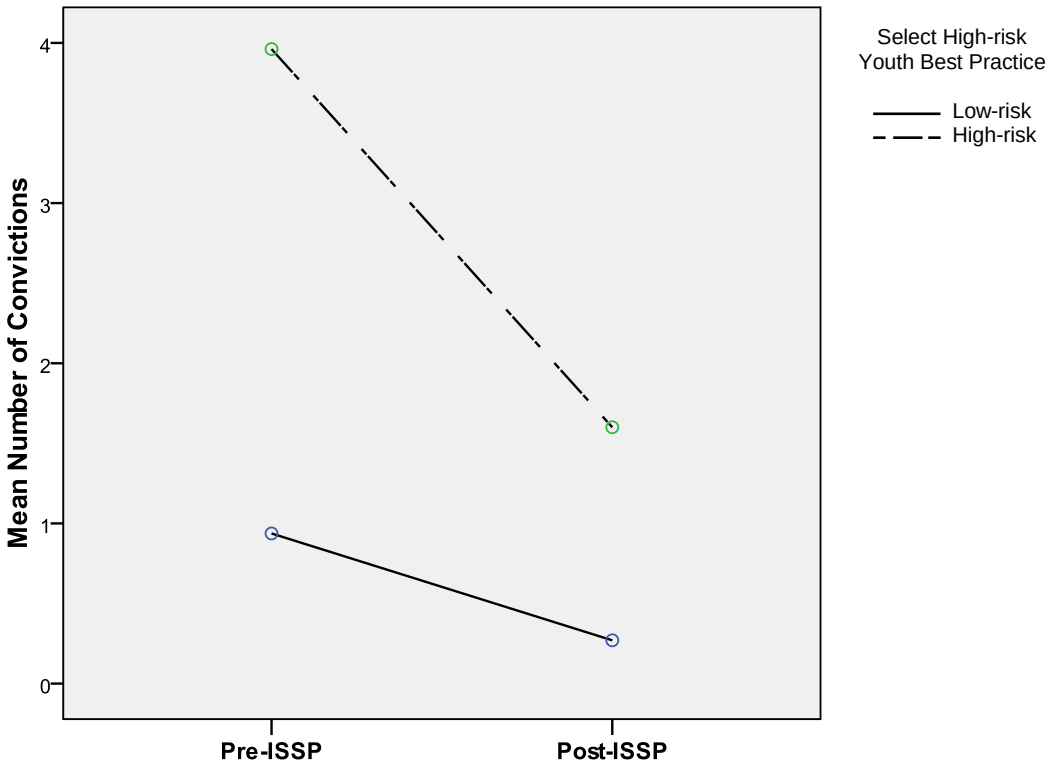
Sample size for addictions $n = 63$. Model -2 Log likelihood = 83.36, $\chi^2(2) = 3.83$, *ns*, Cox and Snell $R^2 = 0.06$, Nagelkerke $R^2 = 0.08$, Hosmer-Lemeshow $\chi^2(8) = 5.03$, $p = 0.76$.

Table 95. Logistic Regression Analysis: Any Recidivism by Engagement

Predictor	*B*	SE(*B*)	Wald χ^2	*df*	*p*	Exp(*B*)	95% CI for Exp (*B*)
Block 1							
No Aboriginal	0.78	0.35	4.90	1	0.03	2.19	1.09 – 4.39
MH issues	0.48	0.47	1.06	1	0.30	1.62	0.69 – 4.04
Low cognitive	0.53	0.43	1.49	1	0.22	1.70	0.73 – 3.94
SAVRY	0.05	0.02	6.20	1	0.01	1.06	1.01 – 1.10
Block 2							
No Aboriginal	0.77	0.36	4.63	1	0.03	2.16	1.07 – 4.38
MH issues	0.46	0.48	0.95	1	0.33	1.59	0.63 – 4.03
Low cognitive	0.70	0.45	2.40	1	0.12	2.01	0.83 – 4.84
SAVRY	0.05	0.02	4.93	1	0.03	1.05	1.01 – 1.10
Engagement	-0.56	0.28	4.18	1	0.04	0.57	0.33 – 0.98

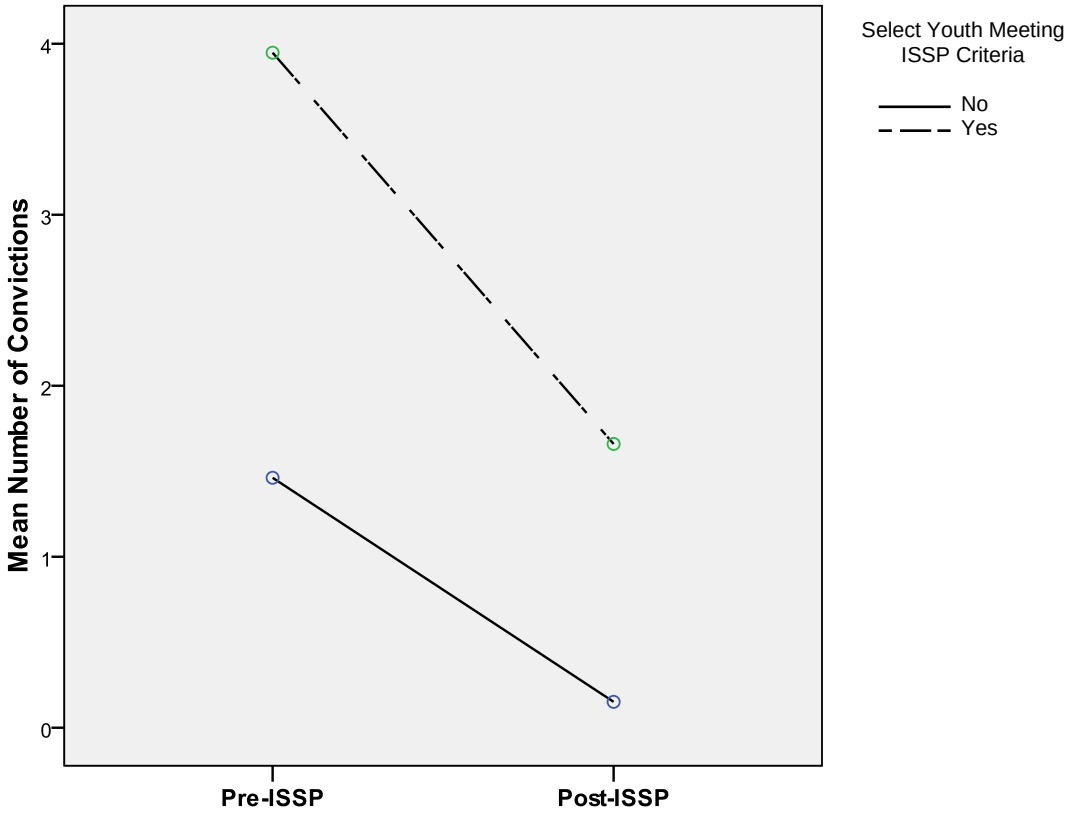
Note. Sample size n = 161. Model -2 Log likelihood = 196.96, $\chi^2(5) = 23.49$, $p < 0.001$, Cox and Snell R^2 = 0.14, Nagelkerke R^2 = 0.18, Hosmer-Lemeshow $\chi^2(8) = 13.19$, $p = 0.10$; Block 2 $\chi^2(1) = 4.34$, $p < 0.05$.

Figure 1. Interaction between Pre- and Post-ISSP Offending and the Selection of High-Risk Youth



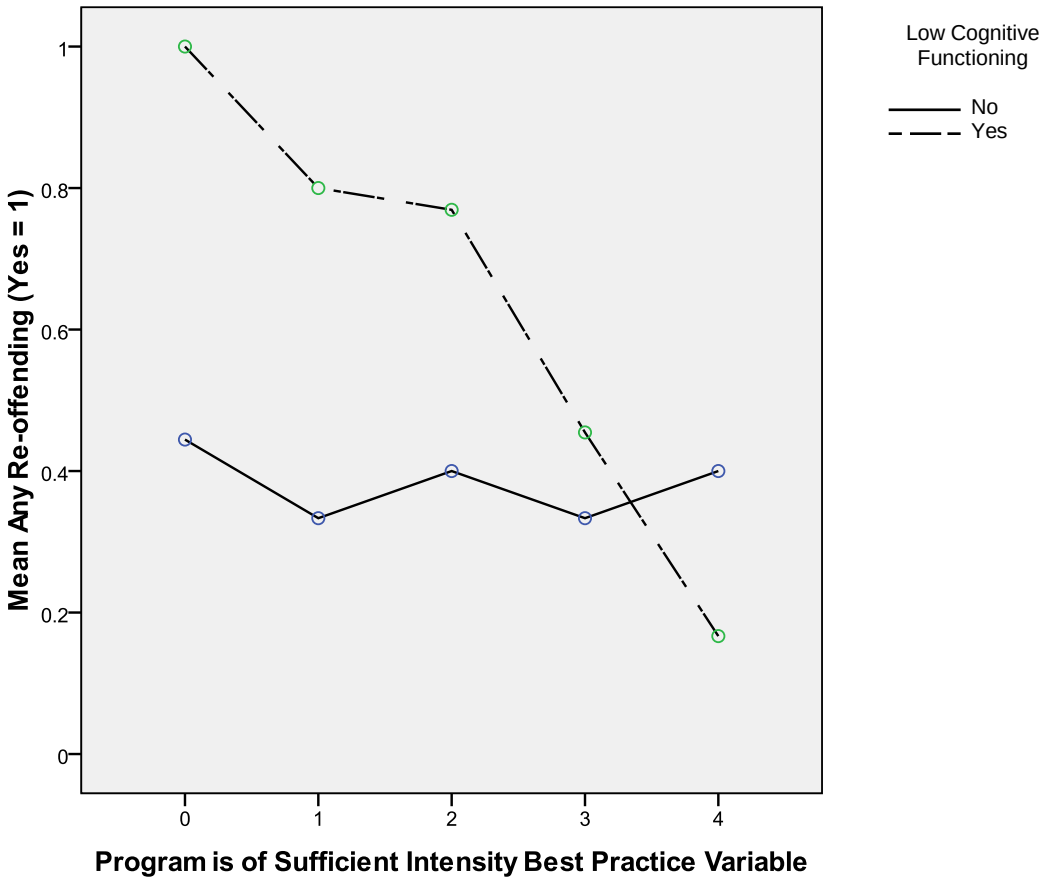
Covariates appearing in the model are evaluated at the following values: Length of ISSP days = 365.95

Figure 2. Interaction between Pre- and Post-ISSP Offending and Fit with ISSP Criteria



Covariates appearing in the model are evaluated at the following values: Length of ISSP days = 360.65

Figure 3. **Interaction between Low Cognitive Functioning and Best Practice Intensity for Recidivism**



Note. See manual for best practice variable intensity values.

Appendix B. Detailed Study Procedures

Due to several issues encountered while reviewing the data, additional decisions and steps were required that are made explicit in this section.

Sources of information. Three sources of information were used in the file-based study. A spreadsheet of Intensive Support and Supervision Program (ISSP) participants was generated and provided by Youth Justice from a pre-existing database. The names of the youth were removed from the file by Youth Justice prior to receipt of the data; case numbers were used to identify youth. The spreadsheet provided by Youth Justice included demographic information such as youth's date of birth, gender, ethnicity, Aboriginal identification, custody centre site from which the referral was accepted, and offences up to March 2012. ISSP variables including start date, end date, and ISSP worker were also supplied by the data file. The spreadsheet was chiefly used to identify youth for the study, run group analyses, and verify the coding of offences.

The youths' paper probation files were the principal source of information for demographic and background variables, Youth Justice ISSP activities, and non-recidivism outcomes. A few probation officers believed that there may be separate ISSP files at the custody centres, but no response was received in a query about these files to the custody centre (personal communication, February 17, 2013). The Youth Justice ISSP guidelines (Ministry of Children and Family Development, 2006) indicate that ISSP information is to be kept on the probation file and moreover, what is available on the file is the documentation that is kept by Youth Justice and would be available to anyone reviewing the program for accreditation and other purposes.

Recidivism was primarily coded from CORNET. CORNET is used to monitor the activities of B.C. Corrections' clients including criminal charges, sentences, and custody or probation status. The risk assessments completed by YPOs using the YCRNA as well as the running notes about the youth's activities from probation and custody are kept on CORNET, although printouts of these are generally also kept on youths' physical files.

Consent. Given that the study used retrospective file data, consent was not sought from study participants. Seeking out participants to obtain their consent could have led to greater risk of harm through the potential for breaches of confidentiality. For similar reasons, parent or guardian consent was not sought. Ethical approval was granted by Simon Fraser University and the project complied with the guidelines for the collection of archival data. Permission to use the data for research purposes was secured through the Youth Justice review process from the Ministry of Children and Family Development. Permission was also secured from B.C. Corrections to access adult recidivism data from CORNET as a number of participants had "aged out" of the youth system during ISSP or the year follow-up period. Furthermore, this access was able to provide official re-offence data for a longer follow-up period, up to the time of data collection.

Sample selection procedures. Youth were selected for the study from the spreadsheet generated by Youth Justice staff. All youth were assigned a participant number to assist with the random selection process. An online random number generator was used to randomly order the participant numbers for the study (random.org, no date). Participants were then selected sequentially from the randomly ordered list of participant numbers for inclusion in the study sample.

Although all participant numbers were included in the random selection process, not all of these youth were eligible for inclusion. Where a non-eligible youth appeared in the randomly ordered list, the reason for non-inclusion was noted and selection proceeded to the next participant. From the information provided by the database, youth who were ineligible due to their ISSP completion date and their age were excluded from the sample immediately after the random ordering phase. The youth who were ineligible due to their non-disclosure period status or location were screened out upon initial search or review of their record in CORNET. The youth with duplicate ISSP services or missing files/file information were screened out as this information became available during the request and review process.

The sample was selected using stratified random sampling to ensure adequate coverage from each custody site. The final proportions did not precisely match those of the total Youth Justice ISSP population due to some files being excluded subsequent to sample selection and these exclusions disproportionately affecting the Burnaby site. In particular, the files that were excluded due to youth receiving duplicate community ISSP services did not appear to be a problem at the other sites. However, the final proportions were roughly equivalent to the original population (Burnaby 61%, Prince George 14%, & Victoria, 25%).

ISSP period selection procedures. Ninety-four of the total Youth Justice ISSP participants had multiple ISSP entries associated with them in the data file. This appeared to occur for at least two reasons. In some cases, youth received two or more distinct periods of ISSP. These were treated as separate programs and the first eligible period of ISSP within the specified time period was chosen if the file was reviewed. As such, youth may have had a previous ISSP period that was completed prior to 2008 and was not included (although previous ISSP was coded as one of the variables; 60 of the 176 youth included in the study had previously participated in ISSP provided by the custody centre or a community provider). Similarly, subsequent distinct ISSP periods were not included. Approximately 46 of the 176 youth included in the study received a subsequent period of ISSP from either a custody or community-contracted provider within the follow-up period.

File numbers also had multiple ISSP entries in some instances when a new ISSP worker was assigned to the youth. In this case, the second ISSP entry was considered a separate period if more than a month separated the completed date of the first period and the start date of the second entry or if the second ISSP entry had the same ISSP worker assigned. Otherwise a change in ISSP

worker was coded and the program was considered to be a single continuous ISSP.

File retrieval process. Files that were selected were pulled by youth probation administrative staff. Both open and closed files that could be accessed in Metro Vancouver were reviewed at the youth probation offices where they were located to minimize the burden on staff. Closed files from sites that were not accessible to the researcher or that were archived at secure storage were sent via inter-office mail for review on site at a Ministry of Children and Family Development office. Unfortunately, due to mobility constraints, open files outside of the Greater Vancouver Area could not be reviewed in the current study.

Pilot file review and development of the data collection sheet. Two pilot files were selected from a youth probation office to help develop and test a data collection sheet created specifically for this study. The variables selected for the study were based on a literature review of variables associated with offending behaviour as well as other negative outcomes and the availability of the information in the probation file. The coding sheet was used to record demographic, risk assessment (YCRNA), ISSP program, and selected outcome variables from the probation files. While some variables were recorded directly from the file, others were coded prior to being entered into the data sheet. Demographic variables recorded directly from the database included age, gender, ethnicity, and site. Background variables recorded from probation files included school status at the start of ISSP and employment status at the start of ISSP as well as indications of FASD, low cognitive functioning, mental health issues, and addictions identified by the youth probation officer.

Risk and needs assessment ratings and other data were recorded from the YCRNA including the completion date of the tool, the name of the youth probation officer who completed it, individual risk and need items, and overall risk, needs, and supervision ratings for all of these on the file. The data coding sheet also included a separate section where risk ratings for the SAVRY and the YLS/CMI 2.0 tools could be entered.

ISSP program information recorded included referral date and information, and needs, goals, contact hours, and summaries from the monthly update logs. The complete ISSP logs were transcribed with identifying information removed in order to provide richer information for the program coding, for verification of quantitative data (e.g., number of contacts), and to allow for the coding of qualitative data. Furthermore, all notes made by youth probation officers about ISSP in their running notes were recorded. Best practice information was generally coded into 5-point variables as described later in this section; however, in some cases where the best practices reduced the amount of variation for continuous data, these variables were also included (e.g., percentage of needs, number of hours, number of months, etc.).

Information on outcomes was collected from probation files. These were normally coded into categories rather than recording each event due to the inconsistency of the format and amount of information across files and to preserve reliability. Variables were generally coded using the categories of low,

moderate, and high, which were provided with specific operational definitions on the coding sheet. Outcomes included legal variables (e.g., failure to report, breaches, arrests, unofficial offending detected by the probation officer), program variables (e.g., number and types of programs the youth participated in during ISSP, attendance, and engagement), as well as other outcomes (e.g., school or vocational program enrolment, employment, periods of unknown whereabouts).

A separate sheet for official legal variables was created for confidentiality purposes and to ensure that raters remained blind to recidivism outcomes when coding file and risk data. Data were recorded from CORNET on prior offending, such as dates, number, and types of charges and convictions, time in custody, and sentence length and types. This included offences which led to the youth's referral to ISSP. Recidivism outcomes for the ISSP period and the year following ISSP completion were also collected, such as number, date, and type of charges and convictions, sentence length and types, time in custody, and offence dates.

File review procedures. Once probation files were randomly selected, individual files were reviewed chronologically and data were collected using the sheets described above. Material placed on the file prior to the commencement of ISSP was reviewed and the risk assessment instruments coded first, in order to minimize potential bias of the risk estimates that might arise if the outcomes were known and to provide an index of the youth's risk prior to receiving ISSP. Although it was impossible to remove all information that would provide an indication of recidivism outcomes prior to coding (e.g., files with several volumes generally indicative of greater justice system involvement), it was not feasible to have youth justice personnel separate out this information prior to review.

Outcome coding from the probation files and the recording of ISSP data were completed as a next step. ISSP program coding was not completed directly from probation files due to concerns that knowledge of outcomes could bias the ratings. Sufficient information was recorded from the files such that the program coding could be completed later. Though the coding of the probation file outcomes would ideally have been completed by a separate rater that was blind to the risk ratings obtained for the file, it was unavoidable due to insufficient resources and the burden of having administrative staff retrieve files multiple times. Measures that were taken to minimize bias included the use of explicit, count-based coding criteria that reduced the amount of subjectivity required. Furthermore, totals for the risk tool were not calculated at the time of coding the files such that raters were not aware of the risk category in which youth fell.

Defining the program outcomes in broad terms was relatively straightforward (e.g., does the program reduce recidivism?), although operationalizing these outcomes was more complex. Recidivism indicators include self-report of offending or official records of further arrests, charges, convictions, and return to custody. Since much criminal behaviour could go undetected by police, self-report is believed to yield the most accurate estimates of true offending but can be fraught with its own problems (see Krohn, Thornberry, Gibson, & Baldwin, 2010) and is rarely available in retrospective studies. Official outcomes tend to be more readily available but there may be great variation among the results for the type of indicator. It is important to

consider that every official measure of recidivism reflects not only the youth's behaviour but also the decision of at least one justice personnel, such that rearrest rates will be higher than for the filing of charges and so on (Harris, Lockwood, & Mengers, 2009). Moreover, charges or convictions resulting from breaches of supervision conditions pose a challenge to the definition of recidivism as the behaviours in themselves may not generally be considered criminal (e.g., not attending an appointment, failing to attend school, or associating with criminal peers; Harris et al., 2009).

Given the procedural and definitional issues with official data, the Council of Juvenile Correctional Administrators (CJCA) issued a white paper to inform the standardization of recidivism outcomes (Harris et al., 2009). Although these are aimed at juvenile correctional agencies to standardize national reporting of recidivism, they can also be helpful in defining program outcomes. The CJCA committee recommended that all studies at least include a measure of adjudication or conviction to minimize false positives and that more than one recidivism measure be included. Another recommendation was to collect information on other factors that might impact an individual's likelihood of justice system involvement, including age, gender, ethnicity, risk assessment ratings, and special needs (e.g., mental health, substance use, low cognitive functioning), due to the latter's disproportionately high probability of re-arrest. To ensure that data are complete and comparable, the CJCA committee's recommendations included separating technical violations from new offences, collecting recidivism data for at least 24 months after the date of interest (e.g., program start), obtaining adult offence data during follow up, and accounting for time at risk. Where possible, these recommendations were followed with some exceptions. The primary dichotomous re-offence variable used throughout the study included both technical violations and new offences since, despite the fact that behaviour leading to breach violations may not be otherwise considered criminal, these outcomes are nonetheless important in that they often extend youths' involvement in the criminal justice system and reflect negative legal outcomes. A separate analysis was run for the overall implementation composite for new offences only so the information would be comparable to other studies. It was not always possible to have a two year period following the start of ISSP due to constraints of the study dates, but these were maximized. Time at risk was accounted for using consistent outcome periods and adult offences were included in the study.

Official offence data were recorded from the database generated by Youth Justice and from CORNET after completing the review of the file; generally a month or more later. This was again done to minimize any bias that might have occurred in the coding due to knowledge of recidivism outcomes. The data were recorded on a separate data sheet as detailed above, such that none of the probation file or program information was available at the time of recording offence data. In addition to the CJCA recommendation, a minimum 18-month period for recidivism outcomes has been suggested for intervention supervision program evaluation, as it was found that most youth who re-offended did so in this time period (Baird, 1991). As such, the current study collected recidivism data up to the period of data collection (August to November 2012). Several measures of recidivism were included to provide more subtle measures of re-

offence than a dichotomous yes/no new offence variable (which was used for the individual best practice analyses), including number of convictions, number of violent convictions, number of indictable convictions, as well as time out of custody. A seriousness index for offending was derived from a 10-point scale for parental offence history from the MacArthur Violence Risk Assessment Study. A composite score was calculated based on the seriousness scores for each of the youths' convictions in the specified time period. The composite provides a rough index of the seriousness of youths' offences over the year following ISSP, although with the caution that the scale was not expressly designed for nor does it appear to have been validated for this purpose.

Data Coding Procedures

In order to reliably code Youth Justice ISSP's program and best practice guidelines within and between raters, it was necessary to create a coding manual. The final version of the coding manual is available upon request.

Program coding. The purpose of the program variables was to calculate an overall fidelity score for the program. These variables were dichotomously coded and defined by empirical research where available, regardless of the distribution of the data in the current sample. Further details on how specific variables were operationalized are included in Chapters 3 to 8.

Variable coding. Individual best practices were coded on a 5-point scale such that analyses could be conducted on the effects of these practices on relevant outcome variables. While theory provided the outline for these variables, considerable modifications were required in the operationalization of the variables to approximate a normal distribution (where possible) according to the characteristics of the current sample. Further details on the specific variables are included in Chapters 3 to 8.

Pilot variable coding procedures. Twenty cases (approximately 10% of study sample) were randomly selected to pilot the coding manual as well as define the normal limits of the variables in the preliminary coding process. These cases were ultimately included in the study.

Once the 20 pilot cases were coded, histograms were created to illustrate the distribution of the data. Modifications were then made to the manual to ensure that every category of the variables meaningfully captured data and to approximate a normal distribution such that the variables could be used in the statistical analyses required for the study. The variables were re-coded according to the new definitions and the distribution analyses were repeated. If the distributions approximated a normal curve at this stage, the next 20 cases were coded; otherwise further modifications were made as needed. Some variables required re-coding several times to achieve a normal distribution. Once the 40 cases were coded, another examination of the distribution was completed and adjustments made as required. The rest of the data were then coded and a final inspection of the distribution of data was made. Three variables required further modifications once coding was complete due to changes in the distribution once the rest of the data were coded.

Appendix C. Excluded Participants

This appendix outlines the issues encountered in selecting youth for the study and the number of youth excluded under each criterion. Analyses of the group differences are outlined and presented in Appendix D.

Ineligible files. Of the 536 total youth participating in Youth Justice ISSP from the year 2008 to 2011, 222 were summarily deemed ineligible due to their failure to meet the study criteria. A total of 125 youth were excluded due to their ISSP end date falling after the eligible cut-off date (September 2011), thus not allowing for a year follow up during the data collection period. A further 97 youth were considered ineligible as their Youth Justice ISSP participation began subsequent to their 18th birthday.

Non-disclosable files. Regarding youth who otherwise met study criteria, the files of 21 youth could not be searched or retrieved at the time of data collection under the period of access outlined in Sec. 119 (2) in the *Youth Criminal Justice Act.* Briefly, the files of youth who have no further involvement in the criminal justice system for a set period of time after they have completed their Court disposition (e.g., end of probation term) cannot be disclosed for research or other purposes. The length of the period depends on the disposition but ranges from three years from a finding of guilt for a conditional discharge order to five years after the completed disposition for an indictable offence. If an adult conviction is acquired during the period, youth convictions remain on the criminal record.

While this factor could serve to bias the results of the study, it was unavoidable due to the legal restrictions around access for youth who have offended. Also, given the three-year period of inclusion for the study, it was still possible to include youth who had offended more recently and who may eventually meet the non-disclosure criteria. Caution is nonetheless warranted in interpreting the results of the study, as it is unknown whether youth who managed to refrain from offending during the disclosure period were more successful ISSP participants, lower-risk youth, or impacted by other factors not captured in the study. It is furthermore unknown how many youth may have been excluded from the initial list sent for this reason, although given that the period of access is generally a minimum of three years from the time of the completion of a disposition (as few ISSP participants only received a conditional discharge), the exclusions were anticipated to be minimal.

Other excluded files. In terms of other eligible youth, 17 more files could not be retrieved as youth continued to have ongoing justice system involvement at the time of data collection and the files were located at a site that was inaccessible to the researcher. Subsequent to the start of data collection, 14 files could not be retrieved by staff or were missing a critical volume of the file. Six more files were identified as having received duplicate ISSP services with a community provider simultaneously for part of or the entire program during the

file or ISSP coding process, such that the effects of Youth Justice ISSP could not be disentangled.

The resulting sample of eligible and presumably available files was 256. Of these, 80 files were not selected for inclusion in the study through the random selection procedure. However, it is possible that some of these may also have been determined to be ineligible or unavailable for review in the course of data collection. Thus the final sample of reviewed files represented approximately two-thirds to three-quarters of eligible youth.

Appendix D. Data Quality Measures

The focus in this section is on information that was missing, as this was the predominant concern with the data; however, there is a brief summary of the inadequacies and apparent inaccuracies in the data included at the end. Several measures were taken to compensate for problems with the data, which are described below. These measures are organized by type and source of information after a brief discussion of issues associated with missing data.

Missing Data Overview

Missing data issues can be endemic to a retrospective file study due to the data required for certain research questions not being included, not being reliably completed, or not being in a usable format. The presence of missing data can pose significant threats to validity for file-based studies depending on how they are interpreted (Långström et al., 1999). The measures employed to minimize these threats are dependent on the types of missing data and the mechanisms by which they are missing. Missing data may be classified under three general types; cases, variables and occasions (McKnight, McKnight, Sidani, & Figueredo, 2007). In the current study, missing cases would include those files that were excluded for unplanned reasons throughout the sample selection process, although technically missing cases are those for which there are no data and it was possible to obtain some information on these cases through the database. Variables for the current study refer to those cases where file data were collected but specific variables were missing information for various reasons. Occasions as a source of missing data also applied to the current study, in that there were variables for which file data were available for multiple time points.

Missing data are furthermore categorized according to three missing data mechanisms in the research literature (Enders, 2010). Data that are missing completely at random (MCAR) are those that are missing in a haphazard fashion and are not related to the data, as in cases where program coding ratings are missing due to the ISSP referral form missing from the file. Missing at random (MAR) refers to data for which the probability that they are missing is related to some other variable in the model of analysis but not the values of the variable with the missing data itself after the relationship with these other variables has been partialled out. For example, specific risk tool ratings may be systematically missing for youth not enrolled in school or for youth with less complete file information. For data that are missing not at random (MNAR), the probability that the data are missing are related to the values of the variable itself, such as the likelihood that re-offence data are missing for youth who completed their disclosure period without reoffending such that the probability of missing data is systematically related to zero values for reoffending. Unfortunately, it is difficult to know for certain whether missing data are random or non-random since making the distinction would require knowing the missing values. Knowing the mechanism by which data are missing has important implications for the

measures selected to compensate for missing data, as many of the commonly used missing data methods (e.g., listwise or pairwise deletion, mean imputation) rely on an MCAR mechanism and are biased under MAR or MNAR (Enders, 2010). Maximum likelihood estimation and multiple imputation procedures assume that data are MAR.

In choosing missing data handling methods, McKnight and colleagues (2007) recommend a decision-making process involving multiple steps and considerations that fall outside the scope of this study. However, depending on the type of missing data, those that are relevant to the current study are to identify the relevant variables, to specify the level of analysis from which data are missing (e.g., individual variables versus composites), determine the missing data features (mechanism, pattern, and amount), and to engage in decision-making according to the relevant objectives. In making these decisions, it is also important to consider the potential consequences of missing data and problems with various handling methods, which could include threats to the internal validity of the study due to bias introduced by systematic missing data, loss of statistical power, and problems with generalizability (McKnight et al., 2007; Roth, 1994).

The missing data handling methods discussed below pertain primarily to the outcome analyses. Since implementation is focused on the actual delivery of the program, methods to replace or impute missing data are inappropriate. Missing data and the sample sizes used in percentage calculations are thus reported as well as the reasons that data were missing.

Missing Cases

A sizeable proportion of cases from the initial list of youth who participated in ISSP during the study period were missing from the final sample for a number of reasons. Of these, 41% of the complete sample of youth did not meet study criteria. For this reason, these "missing" cases do not have any implications for the validity of the data and so further analyses were not conducted. However, it should be noted that systematically excluding these cases may limit the generalization of the study results to youth under the age of 18 and the results may not be fully applicable to the program as it is currently implemented.

Files that met study criteria but were otherwise excluded or constituted missing cases were divided into those that were believed to be excluded by random means (MCAR; 3% of complete sample), those that were excluded due to non-random factors that were not believed to be related to important variables in the study (i.e., those associated with implementation or outcomes, which were effectively MAR; 18% of complete sample), and those that were known to be related to these variables (MNAR, 5%[2] of complete sample). The MCAR group

[2] It is possible that this number was higher as some of the randomly excluded youth may have been found to be non-disclosable upon searching CORNET or to have duplicate ISSP services; however, these are the known MNAR youth

were those files that were genuinely missing, as in the event of a missing file volume. The MAR group were those youth that were excluded due to the location of the open file and those that were not randomly selected, since the selection procedure relied upon stratified sampling and was thus not entirely random. Although the latter are not missing in the sense that the current study was not intended to be based on the entire population of eligible ISSP youth during the study period, it is still beneficial to determine whether these youth were different than those included in the study. The MNAR group were those files that were missing due to youth successfully completing their non-disclosure period and youth that were excluded for receiving duplicate ISSP services, as their likelihood of being excluded was related to their offence outcomes or the implementation factor of intensity of service, respectively.

Demographic and offence information of these missing cases was fortunately obtained at the beginning of the study, which is often not available for missing cases. On the other hand, these data were not sufficiently complete or appropriate for any estimation procedures that might be used for the rest of the data missing from these cases (e.g., ISSP implementation data). Nevertheless, the available data may be used for comparative analysis to assess whether files were missing due to random or "ignorable" factors or whether there were other systematic differences that might threaten the validity of the results or limit the generalizability of the data.

Each of the three groups was analysed individually against the group of youth included in the study to identify whether there were any systematic differences for any of the available data. Pearson's chi-square tests were used to analyse differences for categorical variables and independent samples *t*-tests were used for age and offence variables. Results are presented in Table 1 and 2. There were no significant differences between the included sample and the MCAR and MAR groups in age, gender, Aboriginal identification, and the number of offences, number of violent offences, number of non-breach offences, and instances of offending prior to ISSP and after the start of ISSP. The MAR files were less likely than the study group to include youth from the Prince George catchment region, χ^2 (2, n = 273) = 7.96, $p < 0.05$, which was anticipated based on the study's stratified sampling procedure. Gender and Aboriginal identification approached significance, with the MAR group including marginally more males and more Aboriginal youth.

For the MNAR group, there were no significant differences for age, gender, or Aboriginal identification, although the MNAR group included marginally more males than the study group. There were significantly more youth excluded from the Burnaby catchment region than the other regions, χ^2 (2, n = 203) = 6.77, $p < 0.05$. As noted previously, the Burnaby catchment region was disproportionately affected by the duplicate ISSP exclusion criterion since it seems to have been the only site where duplicate community ISSP was available; when these youth were excluded from analyses, there was no difference between the MNAR and included groups in terms of location. The MNAR group had significantly fewer offences than the included group for nearly all of the offence variables after the start of ISSP (see Table 2). Of particular concern to the validity of the study was the possibility that youth who were low

risk may have been systematically excluded by the non-disclosure criterion. In order to determine whether study and non-disclosable youth were at least comparable at the outset of their participation in Youth Justice ISSP, offence-related variables were analysed as a proxy for risk due to their relationship with subsequent offending (Cottle et al., 2001). There were no significant differences for any of the offence-related variables prior to the start of ISSP between the non-disclosable and study youth.

In sum, from the available data, there were few systematic differences between the three groups associated with the hypothesized missing data mechanisms and fewer still had evidence of being associated with implementation or outcome variables that were central to the study. In the case of the MNAR group, although a group of youth that did not re-offend were systematically more likely to be excluded, the most likely result would appear to be the reduction of power and attenuation of the strength of the relationship between implementation and outcomes since it was presumably the more "successful" (due to ISSP or otherwise) individuals that were missing. It is difficult to know for certain that there were no other systematic differences, since further data were not available but it is possible to at least confirm that youth appeared to be equivalent on offence and other dimensions prior to ISSP.

Missing Data at the Variable Level

For those files that were included in the study, information to code variables was lacking from files for a number of reasons; examples of which are detailed below under the sources of data used in the study. In some cases, data were to be entered directly from a specific form and either the form was missing or the item was not completed. Other variables were coded from counts based on notes, where missing data were less obvious. Several strategies were employed including missing data handling methods, qualitative measures, and augmentation through alternate sources of data that are included in the examples below.

Prior to even considering missing variables in the available file information, it should be noted that the problem with relying on probation officers' records for outcomes is the likelihood that they did not detect or were not informed of outcomes. Given that youth may not report substance use, being absent without permission, or failing to attend school for fear of being breached, or may be reluctant to report outcomes such as self-harm, probation officers may not have been aware of some of their behaviour. Some files included notes from collateral contacts regarding attendance at programs and curfew checks, but these were not uniformly present for the files of all youth. A caution for the following section is that there may be data missing from multiple stages in the current study (e.g., youth informing YPO, YPO deciding to record information, absence of file information, failure to record variable during data collection, etc.) for which there may be limited or no information.

File Information. The amount of information available in the files to code specific variables and outcomes varied widely across youth. While information pertaining to demographic variables such as gender and age was available for all

youth, other information was frequently missing or inconsistent. These problems appeared to be less prevalent with legal variables (e.g., breaches submitted, failures to report, early probation termination), than program or psychosocial outcomes. In particular, there was often very little information contained in the files of low-risk youth.

For a number of variables (e.g., enrollment in school at the start of ISSP), only a single case was missing data and the data were presumed to be missing under either an MCAR or MAR mechanism. Analyses used listwise deletion for these cases, as it is acceptable to use deletion methods when there are few (<1%) missing data for a variable under these mechanisms (McKnight et al., 2007). Larger amounts of missing file data chiefly plagued mental health and substance use variables. Some files included forensic mental health assessments or assessments for FASD, but these were not routinely completed for all youth. The most reliably available source of information was the YCRNA, in which there is an additional needs section with items pertaining to mental health, low cognitive functioning, FASD, and addictions. Despite the wider availability of this information, no rating was available for 10% of study youth for mental health difficulties, 15% of youth for cognitive functioning, 20% of youth for FASD, and 9% of youth for addictions.

A further problem was that, while some variables could be recorded directly from forms on the file where available, other variables had to rely on coding events and instances of behaviour from probation notes. In this case, missing data may have also been introduced by a failure to observe instances of outcomes while coding. Variables were thus coded into discrete categories rather than recording their numerical values, such that if a rater failed to record an occurrence of an outcome or an outcome was not documented in the file, the item would be more likely to retain the same value. Missed data may also cause reliability to be poor for a variable, which was another reason why only variables with good to excellent reliability were included for analysis.

Generally speaking, for mental health variables and outcome data, the default assumption for the file information was that anything not recorded did not occur. Although Långström and colleagues (1999) caution that it is often inappropriate to infer the absence of a variable solely due to its failure to be included in file information, they indicated that for some variables it is acceptable to assign a clear “no” rather than code this as missing data (e.g., whether or not an adult offender has biological children). However, no further guidance was provided on how to determine what variables would meet this criterion. As such, variables were coded as absent where there was a reasonable expectation that such information would be included in a probation file. For example, since it appeared that probation officers generally were diligent about recording outcomes and rarely commented on the absence of outcomes, a lack of documentation on an outcome (e.g., curfew, periods of unknown whereabouts) was assumed to reflect the non-occurrence of behaviour. Similarly, since probation officers are expected to complete YCRNAs, the failure to indicate mental health or addictions problems on this form was interpreted as the absence of these variables. Although it is also possible that probation officers did not feel they had the appropriate training to make such classifications, it was

presumed that failure to note these issues might reflect less severe forms of disorders even if they were present.

Another difficulty with coding the outcomes from the probation files was the variable length of follow-up data available. Some youth finished their probation period immediately after their participation in ISSP (36%) or inside of the year follow-up period (27%). In these cases, unless youth committed another offence and returned to the justice system, file information for outcomes was not available for the time that youth were not under supervision. Since missing data techniques are not appropriate to estimate this form of missing information, the length of follow up was recorded so that it could be included in analyses of the outcome variables. In addition, the quality of outcome information was coded based on both the amount of information available and the length of follow up available so that it could be modelled for in the relevant analyses.

Risk Ratings. It is recommended that the YLS/CMI 2.0 and SAVRY be rated using both file and interview data, as these were the procedures used in validation studies (Borum et al., 2009; Hoge, 2005). However, in the absence of interview data, previous studies have demonstrated that the YLS/CMI and SAVRY or their scales are able to predict violence and offending when rated with comprehensive file information (Catchpole & Gretton, 2003; Gammelgård, Koivisto, Eronen, & Kaltiala-Heino, 2008; Lodewijks, de Ruiter, & Doreleijers, 2008; Marczyk, Heilbrun, Lander, & DeMatteo, 2003; Viljoen et al., 2009; although see Burl, 2012).

The amount of information available for making risk ratings varied based on the youths' level of justice system involvement prior to ISSP. While some youth either had a long history of justice system involvement before being referred to ISSP or may have participated in an ISSP term previously that fell outside of the study period, others were immediately referred to ISSP and had less information on which to base risk ratings prior to ISSP referral. A variable was created in order to control for the amount of information on file, which ranged from a single probation presentence report and a YCRNA to comprehensive forensic reports and program summaries. Unfortunately, the variable did not have sufficient inter-rater reliability to be included in the study (see Appendix E). Although generally efforts were made to restrict the review of information to the material available on file prior to ISSP, information on file after ISSP started that applied to the period before ISSP was used to code items that might have otherwise been left blank (e.g., history of trauma only included in a forensic report completed after ISSP started).

There were generally sufficient data on file to complete the risk assessments, although at times ratings had to be made based on limited examples of behaviour, which may overestimate somewhat the prevalence of risk factors. However, 7% of youth were missing at least one SAVRY item in calculating the total and 8% of youth were missing at least one YLS/CMI 2.0 item. The YLS/CMI 2.0 allows for up to four missing items, such that totals could still be calculated for four of these 14 cases, though it should be noted that risk may be underestimated for these youth (Hoge & Andrews, 2011). There was sufficient information to make professional judgment determinations for the

SAVRY for all youth, which, due to its reliance on judgment as opposed to a total score, allows for missing items. Youth with missing total scores could still be classified into YLS/CMI 2.0 risk categories where the risk category would not change regardless of whether the missing items were rated absent or present for the program variables.

A listwise deletion strategy was used for missing risk total scores as it is an acceptable data handling strategy when up to 20% of data are missing through an MCAR mechanism (Roth, 1994). Missing data analyses were conducted to test the possibility of a MAR or MNAR mechanism for risk scores and the small *t*-values and mean differences were not suggestive of these mechanisms. Although a listwise strategy for MCAR data is acceptable, it is not preferred due to the loss of power (Enders, 2010) and so a multiple imputation procedure was considered. A pilot imputation data set was created for the overall implementation model but yielded similar results to the non-imputed data set. It was thus deemed that the potential for error and additional procedures required for a multiple imputation model outweighed the benefits and added unnecessary complexity to the analyses.

CORNET. Some concerns regarding missing data arose in the process of extracting data from CORNET for recidivism outcomes. Unlike the file, complete offence outcome data were available nearly all youth. However, in addition to certain youths' files being unavailable due to non-disclosure policy, specific offences that had fallen outside of the disclosure period for youth included in the study were unavailable on CORNET. Although missing data entries were made to reflect the occurrence of an offence, no further information was available about the type of offence. Where possible, the type of offence was identified through the Youth Justice ISSP database, as the data were compiled a year prior to CORNET data collection and fewer offences had passed the disclosure date at this time. Between the information available through the database and CORNET, offences falling outside of the disclosure period reflected less than 1% of convictions, although they were MNAR.

Program coding and data. Much of the information needed for the ISSP variables or program coding was missing due to inadequacies in the sources of data to address the research questions, which is a separate issue discussed below. However, in some respects, the lack of information created missing data for variables that might be dealt with through missing data procedures. In particular, some of the variables relied on information from the referral form, which was missing from 11% of files. Imputed data were not used for the missing program variables due the relatively small number of missing cases and the problems with using multiple imputation for binary outcomes (Sterne et al., 2009). Although not ideal due to the loss of power, since the data were believed to be missing through an MCAR mechanism and the individual program variables were not central to the study, a listwise deletion strategy was used (Enders, 2010).

Due to missing data for the program variables that comprised it, the composite ISSP scores were missing for 20.5% of youth. Again, a listwise deletion strategy was used for these cases for the reasons discussed above. An

additional index was also computed for the implementation of the program components only (i.e., not including selection or referral elements), as it both was not missing any data points and pertained more directly to the program's delivery.

Missing Occasions

Variables with multiple data points included pre- and post-ISSP measures of school enrollment and employment as well as pre- and post-ISSP offending. However, for the latter variables, missing data was not particularly problematic due to the relatively small amount of missing data (0.6-4%). Due to these relatively small numbers, listwise deletion was considered to be an acceptable missing data handling strategy. Though these data may be missing through an MNAR mechanism, the strategy was selected to provide a more conservative estimate of the effect of implementation variables (Enders, 2010).

Other Data Problems

Inadequate data. Included in this section are situations where it would be inappropriate to use a multiple imputation procedure or other missing data procedures due to the nature of the variables. These included circumstances where entire critical sources of data were missing from the files. Since the fact that they were missing carried important information, it was not appropriate to simply replace the values and a more qualitative problem-solving approach was required. Furthermore, these missing sources caused data to be missing across variables, such that an imputation procedure would be infeasible. Generally speaking, although the file logs were not specifically designed with the goals of this study in mind, the implementation variables were designed to be coded from the available information, provided that a complete account of activities was available.

The most detrimental and unexpected problem with the data was the absence of the monthly program logs staff use to detail their contacts with youth. Unfortunately, in practice, ISSP logs were not routinely completed and 26% of files had no logs, while the average proportion of logs of file was 40% of youths' time in the program. While it is possible that some logs were completed and not placed on the file by probation officers, the completion rate was generally consistent within ISSP workers and, for at least one case, there was email documentation from the probation officer on file requesting the program logs.

Unfortunately, even the presence of logs on file did not guarantee that ISSP coding and program information could be extracted from them. Multiple versions of logs existed that were designed to collect different information. Despite ISSP guidelines detailing both the number of hours and number of contacts per week, only one or the other was required on a single version of the form. Although some ISSP workers included sufficient detail in their additional notes that the number of contacts could be approximated, this rarely occurred in practice. As such, estimates were used to fill in data based on common values noted in file logs that contain both types of information. For example, one hour was estimated per contact noted in the log or probation and vice versa. Though this procedure was likely to overestimate the duration or number of some

contacts, it was also likely to underestimate others and it was hoped that the total number would approximate the actual intensity of contact. Hours were not estimated for other types of contacts due to the likely unreliability of estimates; where ISSP reported both the number of hours and the number of contacts with youth, contacts tended to be between one and 1.5 hours, while the community and family contacts varied in terms of time spent. Where logs were completed, they generally provided information about total hours as well as direct contact hours.

Furthermore, the summary sections of the file logs took on a wide range of formats, but generally were in the form of updates about the youths' functioning and/or of the activities engaged in as part of the program. Despite some files having most or all logs, those that took the form of updates provided very little in the way of information about specific program activities. At best, logs took on a running daily activity format in which hours, type, and number of contact were specified, but many logs fell far short of this. The quality of file logs for coding and the quality of the file more generally for coding (i.e., including probation notes) were recorded as variables (see descriptions in coding manual). These variables fell in a normal distribution, aside from the higher number of files receiving a zero for having no program logs. The variables also took into account the percentage of logs on file.

In order to compensate for the lack of data available through the missing logs or those with formats for which activities could not be coded, the probation officers' running notes detailing ISSP contacts were also recorded. While generally not as comprehensive as ISSP logs, the probation notes were able to supplement, or replace in cases where no logs were on file, the file log information regarding the number and types of contacts. For example, probation officers were generally meticulous in recording the number of probation appointments that ISSP attended with youth, such that these contacts could be coded from probation notes where file logs were missing. Efforts were made to ensure that such information was not counted twice and the final totals were calculated from both sources. As with the ISSP logs, where one direct contact was noted, it was assumed to be equivalent to one hour.

Although the probation files were able to provide valuable supplementary information, they were not without problems either. Probation officers failed to complete a YCRNA prior to referral to ISSP for 31% of cases, making it difficult to rate program variables where the ISSP referral date was unclear. However, in speaking with a probation officer, she informed that the YCRNA remains open on CORNET for editing for a month period, such that a form may be started or nearly complete prior to submitting a referral but the final version of the YCRNA is not submitted until later (V. Pike, personal communication, August 12, 2013). Thus, probation officers were given the benefit of the doubt that they had completed their ratings prior to submission of the referrals as long as they were completed within a month of the referral date. The final number of cases without YCRNAs on file prior to the ISSP referral was 21%, which were treated as missing data above.

Regarding their ISSP running notes, probation officers often indicated what ISSP "will" do or noted that they would ask ISSP to help youth with something without any further notes to confirm that these events had taken place. On a few rare occasions, probation notes indicated that ISSP was unable to complete the task for various reasons, while other times probation officers made notes to indicate that ISSP had completed the anticipated contact. As such, the inconclusive status of these activities necessitated a conservative approach so that the ISSP worker would not be misrepresented to have engaged in more activities with the youth than had actually occurred. The instances where ISSP was anticipated to engage in activities but there was no confirmation were also recorded.

As with the file outcome data, ISSP activities that were not recorded somewhere in the file were assumed to have not occurred. This was particularly problematic in cases where no program logs were on the file and probation officers made few notes about ISSP as the program would receive a low implementation score even though it may have been well implemented (although one might suppose that individuals who were more diligent about completing paperwork may be more likely to adhere to other program guidelines). However, from a program audit or inquest perspective, any information not documented on the file would not be considered to have taken place.

Inaccurate data. The availability of multiple data sources was advantageous in that it was possible to identify some inaccuracies in the data. A primary example of this was for the mental health variables from the YCRNA. It was often unclear what the basis was for these ratings and whether they had been informed by a mental health professional. In order to minimize threats to the validity of the data, files were coded using the best available information on the file. For example, though a probation officer may have identified a youth as having low cognitive functioning, a forensic report on file indicating that the youth had been assessed as having average cognitive abilities could override the probation officer's rating. Instances of overrides were only made when there was clear evidence that the rating was incomplete or incorrect and these were recorded in the coding notes as well as the justification for the override so that these could be revisited if necessary. Also, given the possibility that less severe effects of fetal alcohol exposure may go undetected, a "possibly present/suspected" level of the FASD variable was created, such that a broader range of information could be included if the probation officer did not mark it as present (e.g., if referred for an FASD assessment and no results on file, if clear evidence youth was that exposed to alcohol in utero in a report). While these efforts attempted to balance having this information available for adequate numbers of youth and the accuracy of such information, these variables should be viewed with caution as an index of mental health and are clearly identified as probation officers' file ratings of mental health concerns.

Other inaccuracies appeared to impact the offence outcome data. Although efforts were made to merge the Youth Justice ISSP database and CORNET, there were inconsistencies in terms of type and number of convictions that were particularly problematic for youth with many offences. As this issue would present a systematic bias, these data were not combined. Though there

was more potential for error from coding variables from CORNET, the data coded from CORNET were used in the analyses as the data provided an extra 18 month follow-up period. Furthermore, the database did not include information about charges, actual days spent in custody, etc.

There were also inconsistencies between the Youth Justice ISSP database and probation file for ISSP information. The database had greater than a month discrepancy with file information for the start dates of 29.0% cases and 27.8% of the end dates, while 10.8% of cases had inconsistent start and end dates. Discrepancies between the file and database were up to six months to a year, such that the length of the program would often be incorrect if coded only from the database. The start and end dates of the program were thus estimated from the probation notes if there was more than a month's difference between dates. Furthermore, the incorrect ISSP worker was listed or a change in ISSP worker was not recorded by the database in approximately 30% of cases. Although this variable was not subject to analysis, it nonetheless underlines the problems with relying solely on database information. Finally, approximately 6% of youth listed in the database did not appear to have received any Youth Justice ISSP program at all. To account for these problems, both database and file information were recorded but file information was used in the analyses as the more accurate source of data.

Appendix E. Inter-rater Procedures and Data

The study required a number of different inter-rater analysis procedures due to the types of data, the type of agreement required, and the number of raters.

File variables. Two undergraduate research assistants collected data for approximately 20% of the cases included in the study. They were trained by the author using two sample cases. For both research assistants, the first stage of training involved the review of a case coded in the pilot study with detailed explanations of the items and ratings. The research assistants then each completed a training case that was reviewed and discrepancies discussed. Ten randomly selected probation files were then coded by the research assistants for the purposes of inter-rater analyses. The reason for the small sample was due to the detailed and lengthy coding required (approximately five to seven hours per file), the larger than anticipated final sample, and limited resources, although this resulted in limited power to detect agreement (see precision analyses below).

For the purpose of inter-rater analyses, variables extracted from the file were divided into file variables for inter-rater agreement (where the focus was establishing level of agreement for pre-defined variables) and individual variables for which it was necessary to establish reliability, including data quality and outcome variables due to their importance for the study. These included dichotomous, ordinal, and continuous variables, which were analysed separately. Missing data were assumed to be MCAR and were removed using listwise deletion, since most of the analyses had adequate power with the variables removed.

Risk ratings. The two undergraduate research assistants noted above also completed risk ratings for the YLS/CMI 2.0 and the SAVRY for each of the training cases and files that they coded in the previous section. Both raters were trained by Simon Fraser University faculty and research staff on the use of these tools with training cases and discussion of gold standard ratings of cases, and had used them in previous research studies of adolescent offenders. The research assistants were provided with the manuals while completing their ratings. As with the file and outcome variables, discrepancies between risk ratings on the training cases were reviewed and discussed.

Other studies have similarly relied on file only ratings of these risk assessment instruments and are able to provide estimates of anticipated reliability coefficients for precision analyses. The SAVRY has yielded reliability estimates for file raters in the excellent (0.77-0.91) range for both the summary judgement ratings and total scores (Burl, 2012; Catchpole & Gretton, 2003; Lodewijks et al., 2008). The YLS/CMI total score has also shown excellent (0.80-0.90) reliability among raters for file-based studies (Catchpole & Gretton, 2003; Marczyk et al., 2003; Olver et al., 2012; Viljoen et al., 2009)

Program coding. A graduate research assistant coded the program variables ascertaining fidelity to Youth Justice ISSP guidelines and best practices. Training involved a detailed review of the coding manual as well as the review of the ratings of two completed cases. A further two participants were randomly selected from the pilot coding cases as training cases, for which the ratings were reviewed and discrepancies discussed prior to continuing with further coding. The rest of the 17 pilot cases were provided to the research assistant for review to provide further scoring examples and an illustration of the distribution of scores.

Twenty cases were then selected randomly from the complete sample based on precision analyses (see results below) to provide the inter-rater data. These cases were coded and inter-rater reliability analyses conducted. Since the estimates ranged from 0.41-0.95, which was lower than desired for most variables, further cases were reviewed. Changes were also made to the coding manual in an effort to improve the clarity of the variables. Another twenty cases were selected and the coding repeated to generate the inter-rater estimates.

Inter-rater analyses. Inter-rater estimates were employed for two purposes in the current study. First, inter-rater analyses were used to ensure that data could uniformly extracted from the probation files and to ensure that risk ratings made using the YLS/CMI 2.0 and the SAVRY were consistent, as multiple raters were involved in data collection. The latter were also important in identifying the most reliable risk index among raters for use in the analyses. Second, inter-rater analyses were necessary to evaluate the psychometric quality of the outcome and program coding variables that were operationalized for the current study. The terms "*inter-rater agreement*" (IRA), referring to conformity among ratings, and "*inter-rater reliability*" (IRR), referring to the variables' ability to distinguish between participants, are used to reflect the conceptual differences between these two purposes (Shoukri, 2011). In both cases, agreement among coders is what is measured, but for the latter, an inference of reliability is made (Krippendorf, 2004). Analyses were thus conducted at the rater and variable level, respectively, for these two purposes.

Along with their separate purposes, the inter-rater analyses served multiple functions in the current study. For the file and risk variables, the inter-rater agreement analyses provided an index of agreement for the coders involved in the study for the purpose of interpreting the results. Furthermore, if the rating of file variables were to show low agreement, it is likely that any inter-rater reliability analyses would also be likely to be low. As such, it is possible for the inter-rater agreement analyses to provide a "baseline" estimate of agreement for the purpose of precision analyses and for comparison for the inter-rater reliability analyses. The outcome variable inter-rater analyses were used to establish whether these variables possessed sufficient reliability to be included in the study. The program variable inter-rater analyses required a specific level of reliability prior to the use of these variables in the outcome analyses for the study, but these could also serve to validate the utility of the framework.

Analyses. Intra-class correlations (ICCs) were used for the ordinal and continuous variables in the study because the models used are able to provide

information about both IRA and IRR (LeBreton & Senter, 2008). A two-way random consistency model using a single measures unit of analysis was employed for the program coding variables since a fully-crossed design was used for the two coders and it was desirable to generalize the findings beyond the study (McGraw & Wong, 1996). Given that different coders were used to collect file information used for the IRA and IRR analyses (i.e., for the file, risk, data quality, and outcome variables), a one-way random model was employed using a single measures unit of analysis due to the inability to separate out effects due to judges (Shrout & Fleiss, 1979).

For dichotomous file and outcome variables, Siegel and Castellan's (1998) Kappa (K) was used because the design was not fully crossed for the coders and due to the potential for bias in the marginal distributions among coders (Hallgren, 2013). Although Hallgren notes that Siegel and Castellan's K may be lowered by prevalence effects (e.g., where some items may be represented at unequal rates in this population), which may have been the case for some of the variables in the current study, this was preferable to the inflation of estimates that might be caused by bias. Inter-rater analyses were not conducted for the dichotomous coding of individual program variables since their purpose was to contribute to an overall score, as inter-rater reliability estimates for transformed (i.e., compiled) variables should be conducted in their final form (Hallgren, 2013).

With respect to cut-offs, where required for the program coding and outcome variables, the guidelines for evaluating K or ICCs from Cicchetti (1994) were consulted; that is, a reliability coefficient of 0.60 to 0.74 indicates a "good" level of clinical significance while a reliability coefficient of 0.75 and above indicates an "excellent" level of clinical significance. Elsewhere, cut-offs of K > 0.60 and ICCs > 0.90 have been suggested to retain variables (Långström, Grann, Tengström, Lindholm, Woodhouse, & Kullgren, 1999). However, Långström and colleagues included other variables yielding lower reliability coefficients due to their importance as risk factors in the study of recidivism. Similarly, in the current study, a "good" cut-off was considered sufficient for outcome variables due to the importance of including them in the study and since the small inter-rater sample size and use of only two coders was likely to lead to lower reliability estimates. Due to the centrality of the program coding variables to the current study and the larger sample size, an excellent level of reliability was desired for these variables.

Results

Precision analyses. Due to the limited resources available for the current study for inter-rater analyses, precision analyses were conducted to determine whether the sample size available would be sufficient to obtain the specified level of precision for the ICCs under a one-way model. For the agreement analyses, multiple observations were able to increase the precision of these estimates; however, for the outcome analyses, the observations were limited to 10 since the variables were the focus of the inter-rater analyses. A planning reliability value of 0.8 was used somewhat arbitrarily since the outcome variable ratings were more subjective than the ordinal variables used in the inter-

rater agreement analyses. Since the current study provided observations for two raters for ICCs, the number of observations required under a one-way model for a 95% confidence interval with the lower limit falling at or above 0.6 was approximately 40 (Walter, Eliasziw, & Donner, 1998). Thus, the number of observations provided may not allow for the level of precision required for the study. For these variables, the confidence intervals were reported (Hallgren, 2012) and corrected for the small sample size.

For the dichotomous outcome variables, the anticipated reliability estimate was derived from the estimate for the dichotomous variables from the IRA analyses. A minimum cut-off of 0.60 was used for the low end of the 95% confidence interval with 80% probability. These figures yielded a sample size of between 16 and 218, depending of the probability of the various outcome events being rated present (Shoukri, Asyali, & Donner, 2004). As such, particularly for events with lower prevalence levels, the number of observations may not have been sufficient to ensure that reliability values under 0.60 were not included in the 95% confidence intervals and these were reported.

For the program coding variables, precision analyses were conducted a priori to determine the number of cases that would be required to provide adequate precision for establishing that ICCs exceeded the desired cut-off with sufficient confidence. The planning value was derived from test-retest reliability estimates from coding completed approximately six months after the initial coding, since the error introduced by individual raters could be modelled for in a two-way random effects model. Inter-rater reliability $ICC_{(2,1)}$ values ranged from 0.79-1.0 for the individual variables across time. The average of the 11 data quality and coding variable values (0.91) was used as a planning value to provide a rough estimate of the number of cases required to establish adequate precision (six variables were excluded from the test-retest estimates due to significant changes made to the coding criteria for the variables in the interim or because they were not included in the original coding). The number of cases required to obtain a reliability coefficient with two raters for which the lower end of the confidence interval would be 0.75 was thus approximately 20 (Bonnett, 2002).

IRA analyses. Separate analyses were run for the dichotomous, ordinal, and continuous variables in the file, which were further grouped according to scale to ensure that variance estimates were unbiased. The actual number of observations for raters thus varied from 10 to 725, due to the nature of the variables included in the file.

For the dichotomous (e.g., absent/present) variables, there was an excellent level of agreement, $K = 0.75$, ($n = 66$; 95% CI = 0.56-0.94; $p < 0.001$). A one-way random effects single measures model yielded reliability estimates of $ICC_{(1,1)} = 0.92$ ($n = 102$; 95% CI=0.88-0.94, $p < 0.001$) and 0.98 ($n = 725$; 95% CI=0.97-0.98, $p < 0.001$) for ordinal variables with three and five levels, respectively. For the continuous variables in the study, $ICC_{(1,1)}$ values for 10 observations were 0.82 (95% CI= 0.46-0.95, $p < 0.01$) for months of follow up and 0.97 (95% CI= 0.88-0.99, $p < 0.001$) for number of logs on the file.

IRA analyses for risk ratings. Inter-rater agreement for the SAVRY total, SAVRY professional judgment, and YLS/CMI 2.0 total was assessed using a one-way single measures random effects model. $ICC_{(1,1)}$ values were 0.75 (n = 7; 95% CI = 0.15-0.95, $p < 0.05$), 0.52 (n = 10; 95% CI = -0.12-0.86, *ns*), and 0.79 (n = 8; 95% CI = 0.30-0.95, $p < 0.01$), respectively.

IRR analyses for data quality variables. A one-way single measures random effects model was used to calculate the ICCs for the data quality variables for the study. Regarding variables related to the quality of the file, the $ICC_{(1,1)}$ for the quality of the file for rating outcomes for all 10 files was 0.63 (95% CI = 0.08-0.89, $p < 0.05$) while the quality of the file for rating risk instruments was 0.40 (95% CI = -0.24-0.80, *ns*). Thus the latter was not used in any of the analyses.

IRR analyses for outcome variables. Inter-rater reliability for the ordinal outcome file variables was calculated using a one-way single measures random effects model. The inter-rater reliability estimates and their confidence intervals are reported in Table 4, corrected for the small sample sizes. The $ICC_{(1,1)}$ values for failure to report to probation appointments, documented breaches of conditions, police contact, incidences of violence, curfew violations, periods of unknown whereabouts, drug and alcohol use, self-harm and suicide attempts, number of other programs, maintenance of employment during ISSP and follow up, form of employment at the end of ISSP, enrollment in school throughout ISSP and follow up, ISSP attendance at integrated case management meetings, and probation officers' rating of youths' ISSP engagement exceeded the a priori stipulated cut-off of at least good clinical significance. Variables that did not meet the cut-off criteria were hospitalizations, probation appointments attended by ISSP, file documentation of arrests, and informal documentation of offence behaviour (i.e., for suspected or confirmed involved in criminal activity for which charges may not have been submitted).

For the dichotomous outcome variables, inter-rater reliability was calculated using Siegel and Castellan's K. Variables indicating whether youth held any full-time employment during ISSP and follow up, whether youth attended school at all during ISSP and follow up, and whether or not youth were enrolled in school and employed at the last available follow up all exceeded the cut-off of 0.60. The variable rating of whether youth held any part-time employment during ISSP and follow up fell short of the cut-off.

Furthermore, the author additionally had access to the CORNET file notes for all cases and all cases were checked for consistency purposes. Adjustments were made where there was clear evidence in the notes that contradicted a rating, which may have occurred due to the CORNET printouts missing from the physical file, for example.

IRR analyses for program coding variables. Inter-rater reliability for the program coding variables was calculated using a two-way single measures random effects model. For clarity, the IRR analyses for the individual variables associated with each best practice are included in the specific results sections associated with the variables (see also Table 5).