

Bridging Risk Assessments to Case Planning:
Development and Evaluation of an Intervention-Planning Tool for Adolescents on Probation

Jodi L. Viljoen Dana M. Cochrane Catherine S. Shaffer Nicole M. Muir
Simon Fraser University

Etta Brodersen
Nova Scotia Health Authority

Billie Joe Rogers
Reciprocal Consulting

Kevin S. Douglas Ronald Roesch
Simon Fraser University

Robert J. McMahon
Simon Fraser University and B.C. Children's Hospital Research Institute

Gina M. Vincent
University of Massachusetts Medical School

Criminal Justice and Behavior

DOI: <https://doi.org/10.1177/0093854819873019>

Author Note

Jodi L. Viljoen, Dana M. Cochrane, Catherine S. Shaffer, Nicole M. Muir, Kevin S. Douglas, and Ronald Roesch, Department of Psychology, Simon Fraser University. Etta M. Brodersen, Nova Scotia Health Authority. Billie Joe Rogers, Reciprocal Consulting. Robert J. McMahon, Simon Fraser University and B.C. Children's Hospital Research Institute. Gina M. Vincent, University of Massachusetts Medical School. The members of the ARROW international advisory board include (in alphabetical order): Drs. Keith Cruise, Kevin Douglas, Adelle Forth, Laura Guy, Ed Hilterman, Robert McMahon, Marlene Moretti, Ronald Roesch, Gina Vincent, and Billie Joe Rogers.

This research was supported by a grant from the Social Sciences and Humanities Research Council of Canada. The authors would like to thank the many people who shared their expertise and guidance, including Indigenous advisors, the youth justice practitioner committee, and the international advisory board.

Correspondence concerning this article should be address to Jodi Viljoen, Department of Psychology, Simon Fraser University, Burnaby, BC V5A 1S6. Contact: jviljoen@sfu.ca

Abstract

Even though risk assessment tools are often intended to inform case planning, they do not provide much direct guidance. As such, we developed an intervention-planning tool called the Adolescent Risk Reduction and Resilient Outcomes Work-Plan (ARROW) to accompany the Structured Assessment of Violence Risk in Youth. The ARROW includes a decision support system, guide, and training, and is one of the first tools of its kind. To evaluate the ARROW, we conducted two studies: (1) a vignette study with 178 professionals, and (2) a field study with 320 propensity-score matched adolescents. Most professionals (> 98%) rated the ARROW as useful. Moreover, compared to (a) unstructured plans and (b) a simple form, ARROW plans were more likely to include supported interventions, adhere to best practices, and integrate culturally-tailored approaches for Indigenous adolescents. Formulations also showed improvements. However, further research is needed on strategies to bridge risk assessment and risk management.

Keywords: risk management, risk assessment, case management, formulation, SAVRY, culture, Indigenous

Bridging Risk Assessments to Case Planning: Development and Evaluation of an Intervention-Planning Tool for Adolescents on Probation

Most youth probation agencies in the United States and Canada have adopted risk/needs assessment instruments for improving case management, disposition planning, and treatment (Hannah-Moffat & Maurutto, 2003; Wachter, 2015). Youth probation officers (YPOs) assess youth on their caseload with these assessment instruments, and then afterwards, develop case plans that include treatment recommendations and other risk management strategies (e.g., supervision). These plans serve as the focal point for intervention delivery. However, despite their importance, agencies typically do not provide their staff with much guidance on how to develop these case plans, nor is this information provided within risk assessment instruments themselves. Instead, there appears to be an implicit assumption that after rating risk, professionals will inherently know how to develop appropriate plans.

This assumption, unfortunately, is often incorrect. Even when risk assessment instruments are used, many of adolescents' needs still go unaddressed and untreated (Nelson & Vincent, 2018; Peterson-Badali, Skilling, & Haqanee, 2015; Viljoen, Cochrane, & Jonnson, 2018). Thus, in a prior study, we examined whether the adoption of a simple case planning form, alongside the Structured Assessment of Violence Risk in Youth (SAVRY; Borum, Bartel, & Forth, 2006), resulted in better case plans as compared to standard practices (i.e., an unstructured case plan and a non-validated tool; Viljoen, Shaffer, Muir, Cochrane, & Brodersen, 2019). The results of that study indicated that when YPOs used the case planning form, their plans were more likely to include interventions to address adolescents' needs. Nevertheless, this form was simple; it prompted YPOs to target adolescents' needs (e.g., substance use) but did not provide guidance on which interventions have support (e.g., cognitive-behavioral therapy).

As such, in the current project, we extended this research by testing whether we could further improve case plans by developing and implementing a multicomponent intervention-planning tool called the Adolescent Risk Reduction and Resilient Outcomes Work-Plan (ARROW; Viljoen, Brodersen, Shaffer, Muir, & ARROW Advisory Board, 2014). To evaluate the ARROW we conducted two studies: (1) a vignette study in which YPOs developed case plans with and without the ARROW, and (2) a pre and post field study in which we compared the quality of plans made with the ARROW to those made with a simple case planning form. The purpose of these studies was to provide an initial evaluation of whether the ARROW improves the quality of professionals' case plans, consistent with its purpose. However, we did not examine if improved case plans or adherence to the ARROW leads to better outcomes for youth.

Problems with Case Plans

The primary impetus for the ARROW was that our community partners – a provincial youth justice authority – approached our team and requested assistance in improving case plans. Although this agency used the SAVRY, a well-validated risk assessment tool that includes a focus on modifiable treatment-relevant factors, administrators expressed concern that SAVRY assessments were not optimally being used to develop effective case plans. This problem is not unique to this agency, but is instead, a pervasive problem that many agencies face. Indeed, a growing body of research indicates that professionals often show a lack of follow through on risk

assessments (Bonta, Rugge, Scott, Bourgon, & Yessine, 2008; Viljoen et al., 2018).

Even when professionals use risk assessment tools to identify adolescents' needs, on average, they only address 31% to 63% of these needs in their plans (Bosker, Witteman, & Hermanns, 2013; Dyck, 2016; Luong & Wormith, 2011; Singh et al., 2014; Vieira, Skilling, & Peterson-Badali, 2009). In other words, overall match to the *need* principle of the risk-need-responsivity model (RNR; Bonta & Andrews, 2017) appears modest in most studies. This is problematic because, when most of youths' needs are addressed, youth are less likely to reoffend (Luong & Wormith, 2011; Peterson-Badali et al., 2015). Also, professionals do not necessarily select and implement interventions that have research support. In fact, researchers estimate that only 5% of adolescents with a history of serious offending receive evidence-based treatments (Henggeler & Schoenwald, 2011). Another common problem is that in their plans YPOs often reiterate probation conditions or list services, rather than planning a broader range of strategies that they themselves will use in their sessions. This is problematic because probation officers are not only "service brokers" but also "agents of change" (Bourgon, Gutierrez, & Aston, 2012), and as such, the strategies that they themselves use make an impact. For instance, when probation officers use techniques such as problem solving, prosocial modelling, and reinforcement, clients are less likely to reoffend (Bonta et al., 2011; Chadwick, DeWold, & Serin, 2015; Trotter, 1996). Finally, professionals do not routinely tailor plans to youth's race, ethnicity, or culture. As an example, in Canada, Indigenous people are nearly eight times more likely to be incarcerated than non-Indigenous people (Roberts & Reid, 2017) due to factors such as intergenerational impacts of colonialism (McCaslin & Boyer, 2009; Monchalin, 2016; Shepherd & Ilalio, 2016). However, calls for culturally-informed approaches have remained unmet (Truth and Reconciliation Commission of Canada, 2015).

To help address these problems, our team of researchers, Indigenous advisors, and practitioners developed the ARROW through a community partnership approach. The objectives of the ARROW are to: (1) increase the inclusion of empirically-supported interventions in case plans, (2) expand the breadth of interventions beyond surveillance, (3) increase the inclusion of culturally-tailored strategies, and (4) improve the general quality of plans (e.g., extent to which plans address adolescents' needs).

It is important to emphasize that the ARROW is not a treatment program for youth, nor does it teach YPOs how to be "therapists." Furthermore, simply assessing risk and writing a good case plan is unlikely to reduce reoffending directly. Instead, many other steps must occur, for instance: YPOs must follow through on delivering their case plan; evidence-based programs must be locally available; YPOs and therapists must have adequate skills to effectively deliver interventions; interventions must be evaluated and monitored; youth must be motivated and engaged in interventions; youths' caregivers must provide enough support; and service providers must have adequate systems in place to share information. Instead of offering a supposed panacea to tackle all issues, the ARROW focuses on providing professionals with guidance on how to develop case plans, as the SAVRY does not include this information. In particular, the ARROW aims to help professionals make the best possible decisions about interventions in light of the available options.

Development of the ARROW

Phase 1: Needs Assessment. To develop the ARROW, we consulted with intended users and other stakeholders (i.e., YPOs, social workers, psychologists, managers, administrators; $n = 15$) about what type of case planning approaches might be useful and feasible. For instance, although we discussed the possibility of creating a list of local programs (e.g., a service referral grid), our partners provide services to over 100 communities, and program availability varies by community. Furthermore, programs tend to change over time in regard to staffing, approach, and funding, and as such, a list of programs would quickly become outdated. Instead, our partners indicated that they needed a tool that provided basic education about which types of programs have research support (e.g., cognitive-behavioral therapy), but also incorporated broader intervention strategies that YPOs could feasibly deliver (e.g., motivational interviewing, problem solving).

Phase 2: Consultation with Indigenous Advisors. To help ensure that we attended to cultural considerations, we consulted with Indigenous advisors throughout the province ($n = 26$; authors' names redacted for blind review, 2012). This consultation was led by an Indigenous member of our team and grounded in the 4Rs model, which emphasizes the importance of showing respect for Indigenous peoples, using approaches that are relevant to Indigenous world views, demonstrating reciprocity (e.g., giving back to Indigenous peoples rather than engaging in a one-sided taking of knowledge), and deferring to Indigenous peoples' power to exercise responsibility over their own lives (Kirkness & Barnhardt, 2001). Advisors stressed the importance of including a cultural module within the ARROW and offered ideas for specific interventions strategies for YPOs to use in their work with Indigenous youth.

Phase 3: Literature Reviews. In addition to consultation, we conducted literature reviews to identify evidence-based interventions for each of the 24 risk factors on the SAVRY (e.g., substance use), and each of the six protective factors on the SAVRY (e.g., strong attachments and bonds) (Viljoen, Brodersen, Shaffer, & McMahon, 2016). Consistent with the American Psychological Association's (2006) definition of evidence-based practice, we focused on strategies that are based on the "best available research" (p. 273).

Phase 4: Feedback and Pilot Testing. After developing a draft of the ARROW, we collected feedback from 22 practitioners (e.g., YPOs, psychologists), 5 Indigenous advisors (e.g., an Elder), and an international advisory board consisting of 10 risk assessment and intervention researchers. In addition, we pilot tested the ARROW with 11 YPOs and managers who used the ARROW to develop case plans for youth on their case load ($n = 23$ youth). Finally, we conducted a survey with the broader pool of YPOs ($n = 91$) to check that the intervention strategies in the ARROW (e.g., problem solving) were feasible and that they saw the provision of these strategies as falling within their role as YPOs.

Strategies by Which to Improve Case Plans

One of the primary challenges in developing the ARROW was deciding on the format that it should take. To guide this decision, we reviewed research on implementation science (e.g., Michie, van Stralen, & West, 2011). Implementation science is the scientific study of methods to facilitate the uptake of research findings into real-world practice (Eccles et al., 2009).

Best practice guides. One potential strategy by which to improve case plans is to

develop educational guides that accompany risk assessment tools. For instance, the authors of Historical Clinical Risk-20 (HCR-20; Douglas, Hart, Webster, & Belfrage, 2013) developed a risk management guide, with research-supported interventions, to accompany the HCR-20 (Douglas, Webster, Hart, Eaves, & Ogloff, 2001). Such guides remain scarce, and to date, research has not tested whether they improve case plans. However, although guides may be a useful starting point, educational materials alone typically have only a modest impact on practice, possibly because professionals do not regularly read or refer to guides (Fixsen, Naoom, Blasé, Friedman, & Wallace, 2005; Giguère et al., 2012).

Decision support systems. Whereas guides only work if professionals read them, decision support systems strive to actively structure the decision-making process through forms or protocols that are embedded into professional practice. As an example, if the YPO identifies that a youth has substance use difficulties, an online system could automatically provide YPOs with some recommended interventions. Decision support systems have been found to improve decision making in a variety of healthcare settings (Hunt, Haynes, Hanna, & Smith, 1998; Moja et al., 2014). In addition, several studies have found that such protocols improve probation officers' case plans (Bosker, Witteman, Hermanns, & Heij, 2015; Bosker & Witteman, 2016; Viljoen et al., 2019). For instance, Bosker and Witteman (2016) compared the quality of plans that were developed with a risk assessment tool alone versus a tool *plus* a decision support system. They found that when probation officers used a tool plus a decision support system, their case plans addressed a greater proportion of their clients' needs.

Training. Yet another approach is to provide professionals with training. Again, work in this area is limited. Although the Strategic Training in Community Supervision program (STICS) has gained popularity and achieved positive results (Bonta et al., 2011), this program focuses on teaching probation officers to deliver cognitive-behavioral therapy to address procriminal attitudes, rather than on how to write case plans. Also, even though training might improve case plans, training alone is often insufficient in creating long-term changes in professional practice (Fixsen, Blase, Naoom, & Wallace, 2009).

Case formulation. A final approach is to provide professionals with guidance on case formulation (Douglas et al., 2013; Hart et al., 2003). Case formulation is the “process or product of gathering and integrating diverse information to develop a concise account and nature and etiology” of the factors that contribute to offending (Hart, Sturmey, Logan, & McMurrin, 2011, p. 118). By developing a strong case formulation, probation officers might be better able to develop case plans that target relevant risk factors and plan for scenarios that might occur, but research is limited. Currently, probation officers receive inadequate guidance on formulation (Brown & Völlm, 2013), and most adolescent risk assessment tools, including the SAVRY, do not include instructions for writing formulations.

Although each of the above approaches hold promise, on their own, they may have only a modest impact. As such, we designed the ARROW to be a multicomponent tool that incorporates each of these approaches. To increase its efficiency, the ARROW decision support system automatically populates with ideas for possible interventions rather than focusing on open-ended questions that require considerable documentation. Altogether it takes approximately 20 to 30 minutes to complete a case formulation and plan using the ARROW.

Components of the ARROW

Decision Support System. The ARROW's decision support system guides professionals through the case planning process through a computerized form. The form prompts professionals to select risk and protective factors to include in their plan; based on the factors selected, it populates with suggested interventions. For instance, if the professional checks that they would like to target substance use, the form populates with a list of possible strategies that are ordered from least intensive (e.g., provide caregivers with information on supervision strategies, use motivational interviewing) to most intensive (e.g., refer youth to a supported therapy such as cognitive-behavioral therapy). From this list, professionals can choose strategies that are feasible, commensurate with adolescents' risk level and motivation, and consistent with professionals' own expertise. Professionals can also opt to add-in other strategies if they prefer. In addition, the form prompts professionals to incorporate culturally-tailored approaches, if appropriate, particularly for Indigenous youth. Again, it automatically populates with suggested strategies. For instance, professionals are encouraged to assess cultural connectedness and youths' interest in culturally-tailored services. After professionals have checked off the strategies that they plan to use, the form prompts them to describe how and when they will implement this strategy. Finally, professionals are directed to seek feedback and approvals on their proposed plan from the adolescent, the adolescent's caregivers, and their supervisor.

Best Practice Guide. The ARROW Guide is a 120-page educational guide that accompanies and expands on the decision support system. Part one of the guide, Introduction, consists of instructions on how to use the ARROW to create plans. Part two, Risk Factors, describes supported intervention strategies for each risk factor on the SAVRY (i.e., 2- to 3-pages per factor). These intervention strategies are diverse and encompass not only evidence-based therapies (e.g., Functional Family Therapy; Alexander, Waldron, Robbins, & Neeb, 2013), but also strategies to engage adolescents and their families, build skills, and provide supervision (e.g., motivational interviewing, problem-solving; McCart & Sheidow, 2016; Mahoney & Stattin, 2000; McMurrin, 2009). Part three, Protective Factors, describes supported intervention strategies for each protective factor on the SAVRY. Part four, Diversity, includes a module on culturally-tailored strategies for Indigenous adolescents. It also includes sections on trauma, mental health, and gender. The final section of the guide consists of worksheets and resources, all of which are optional and are intended to provide professionals with ideas and extra resources for case planning. These include example worksheets YPOs can use in their sessions with youth (e.g., Solving Problems Worksheet), or contacts with caregivers (e.g., Caregivers Pamphlet). There are also worksheets that professionals can use to further assist in case planning (e.g., Service Inventory to help offices take stock of local programs).

Training. The ARROW training is two days long, and uses case examples, demonstrations, and role plays to aid in skill acquisition (see Beidas, Cross, & Dorsey, 2014). During Day One, professionals read a case vignette and then write a case plan using standard, unstructured practices. We use this activity to illustrate problems with unstructured case plans, and to highlight characteristics of good case plans. Following this, we provide professionals with an overview of the ARROW, including its rationale and components. During the remainder of Day One and Day Two, we review the basic intervention strategies in the ARROW (e.g., building motivation and rapport). For instance, YPOs role play problem solving with a youth, and watch demonstrations of motivational interviewing. In addition, professionals write a case

plan with the ARROW and obtain general feedback.

Case Formulation. Although the ARROW focuses on how to write case plans, the ARROW's decision support system also provides a template and instructions for writing a case formulation. In addition, approximately two hours of the ARROW training are devoted to formulation and scenario planning. Given that the SAVRY includes 24 risk factors and 6 protective factors, it is not viable to target every factor in an adolescent's case plan. As such, formulation helps YPOs prioritize which factors to include in the plan. However, the ARROW's template for writing a formulation is much shorter and simpler than that used for tools such as the HCR-20 V3 because our partners wanted an approach that was quick and feasible, especially for YPOs who typically have less experience with case formulation than psychologists.

Study 1: Vignette Study

After developing the ARROW, we evaluated it using a case vignette study; this provided an opportunity to efficiently test the ARROW in a controlled setting before our partners implemented it in real-world practice. In our evaluation, we tested whether the ARROW met its primary goal of improving case plans; namely, whether it: (1) increased the inclusion of empirically-supported interventions in case plans, (2) expanded the breadth of interventions beyond surveillance, (3) increased the inclusion of culturally-tailored strategies, and (4) improved the general quality of plans (e.g., extent to which plans addressed adolescents' needs). Given that a secondary goal of the ARROW is to improve formulations, we also examined the ARROW's impact on the quality of formulations (e.g., level of detail).

Method

Participants

The final sample included 178 participants. Approximately half were YPOs ($n = 91$), and the remainder were supervisors or managers ($n = 30$) or service providers ($n = 43$). Half of the participants were female (52.7%, $n = 87$), and most were between 40 and 59 years old (66.4%, $n = 109$). The majority of participants reported that they had post-secondary training (e.g., bachelor's degree; 72.4%, $n = 129$). On average, participants had worked with adolescent offenders for a mean of 15.90 years ($SD = 9.14$).

Procedures

Design. All data collection procedures adhered to ethical guidelines (American Psychological Association, 2017; Canadian Psychological Association, 2017). We used a pre-post study design, wherein professionals conducted an unstructured plan and formulation at the beginning of the ARROW training and an ARROW plan and formulation at the end of this training. Two trainers conducted six training workshops throughout the province using a standardized curriculum (described earlier). One of the trainers was a researcher and registered psychologist, and the other trainer was an Indigenous graduate student in Clinical Child Psychology with 15 years of prior experience as a service provider (e.g., in Indigenous settings). Each workshop was two days long and had approximately 20 to 35 attendees. Although some YPOs missed training due to illness or scheduling conflicts, most YPOs in the province attended (93.5%, $n = 102$), as did managers and local service providers. All attendees were invited to

participate in the evaluation and were provided with an informed consent form. Of the 200 workshop attendees, 94.0% ($n = 188$) chose to participate in the evaluation. However, 10 of these individuals did not hand in any of the evaluation measures and were thus excluded.

Unstructured condition. At the beginning of training, participants were provided with a fictional case example of a 16-year old Indigenous female adolescent (“Anna”) who was placed on probation after being charged with assault. Anna was described as having several risk factors, such as a gang-involved boyfriend and substance use difficulties, as well as protective factors, such as strong connections with family members. Participants were given a two-page description of Anna’s SAVRY assessment, including all item ratings and a written description, and were asked to develop a case formulation and intervention plan. They were provided a blank page to do so and were not given any additional instructions. Participants were already familiar with the SAVRY, as the youth justice agency had adopted the SAVRY two years prior to this time point.

ARROW condition. At the end of the training, participants completed a formulation and intervention plan using the ARROW decision support system (i.e., case planning form). This form provides instructions on how to write a formulation, as well as suggested interventions for each of the SAVRY risk and protective factors. We used the same case example as completed pre-training (the “Anna” case) so that we could more clearly examine the impact of the ARROW. Given that adolescents need different interventions, if we had used separate cases, it would be difficult to ascertain if outcomes were due to the ARROW or differences in the cases. At the end of training, participants completed questionnaires on the perceived acceptability and feasibility of the ARROW (e.g., whether it was useful).

Coding. After the trainings were completed, research assistants (RAs) coded the quality of participants’ formulations and plans. The RAs included three graduate students and three upper-division undergraduate students who completed one day of didactic training and five practice cases. In rating formulations, RAs were kept blind to the study condition. However, it was not possible to keep raters blind for plans, as the ARROW uses a distinctive and recognizable format for plans. To examine interrater reliability a separate rater independently coded a random sample of 48 case formulations and 55 intervention plans.

Missing data. Some participants did not complete and hand in the perceived acceptability and feasibility questionnaires ($n = 9$), the unstructured or ARROW plan ($n = 31$), or the unstructured or ARROW formulations ($n = 44$). Participants with and without missing data did not significantly differ in sex, age group, profession, or years of experience.

Measures

Adherence Checklist for the Evaluation and Reduction of Risk (ACERR). To measure the quality of formulations and case plans, we used the ACERR (Viljoen et al., 2019). This tool was developed through a review of recommended practices (e.g., Bonta & Andrews, 2017), expert feedback, and pilot testing. The Formulation section includes five items (e.g., the formulation describes how risk factors contribute to offending), and the Plan section includes 11 items (e.g., the plan targets key risk factors; see Tables 1 and 2). We omitted three items that were not relevant to the vignette design (e.g., plan was developed in collaboration with the adolescent). Each item was rated on a 4-point Likert scale (0 = *substantial limitations*, 1 = *some*

limitations, 2 = *meets expectations*, 3 = *exceeds expectations*), and then summed. In a prior study, the ACERR was found to have excellent interrater reliability (Holloway, 2015). In the current study, ICCs (two-way random effect model, single raters, absolute agreement; McGraw & Wong, 1996) were .79 for the Formulation Total and .83 for the Plan Total (i.e., “excellent” range; Cicchetti, 1994).

Inventory of Interventions. To examine the types of interventions included in plans, we developed an inventory of interventions through a literature review (Viljoen et al., 2016), and pilot testing. The inventory includes the following sections: (1) Engagement and Skills (e.g., identify triggers; 6 items; ICC = .96); (2) Supervision and Structure (e.g., provide incentives; 5 items; ICC = .84); (3) Families and Supports (e.g., arrange for adolescent to spend time with prosocial supports; 7 items; ICC = .92); (4) Services with Research Support (e.g., refer youth to cognitive-behavioral therapy; 5 items; ICC = .75), (5) Any Services (e.g., refer to job training program; 13 items; ICC = .69), and (6) Culturally-Tailored Strategies (e.g., assess cultural connectedness; 6 items; ICC = .78). Except for the Any Services section, the interventions in this measure are those that have research support or are consistent with recommended practices. Each item was rated on a 3-point Likert scale (0 = *no*, 1 = *possibly*, 2 = *yes*), and total scores for each section were summed. ICCs fell in the excellent range for all sections except for Any Services, which fell in the good range (Cicchetti, 1994).

Perceived Acceptability and Feasibility. Participants rated five items on the ARROW’s acceptability (e.g., How helpful is the ARROW Guide?) using a 4-point Likert scale (*very helpful*, *helpful*, *not that helpful*, *very unhelpful*; see Table 3). In addition, they rated three items pertaining to its feasibility (e.g., The ARROW seems fairly easy to use) on a 5-point Likert scale (*strongly agree*, *agree*, *not sure*, *disagree*, *strongly disagree*). These questions were developed based on a review of similar measures (Proctor et al., 2011).

Data Analysis

To test the impact of the ARROW, we conducted repeated-measure analyses of variance (ANOVAs), using IBM SPSS®, Version 22 (IBM Corporation, 2013). We interpreted η^2 of .01 as small, .06 as moderate, and .14 as large (Cohen, 1988). Given that some outcome variables showed substantial skew, we also conducted a non-parametric alternative, the Wilcoxon signed-rank tests, when skew was greater than 1.0. The effect size for the Wilcoxon test was calculated using the formula in Tomczak and Tomczak (2014). An effect size of .10 is considered small, .30 is moderate, and > .50 is large (Cohen, 1988).

Results

Formulations

Compared to formulations in the unstructured condition, ARROW formulations were rated as significantly higher in quality. More specifically, ARROW formulations scored higher on the ACERR Formulation Total (large effect), as well as on every item (see Table 1). For instance, they were more detailed and specific, and more likely to describe the role of risk and protective factors and types of future offending that might occur.

Plans

Compared to unstructured case plans, ARROW plans scored higher on the ACERR Plan Total (large effect) and on all but one item (see Table 2). In particular, ARROW plans were more detailed, and were more likely to target key risk factors and be tailored to culture. In addition, ARROW plans were more likely to include interventions with support, including strategies related to Engagement and Skills, Supervision and Structure, Families and Supports, Services with Research Support, and Culturally-Tailored Strategies; these effect sizes were large.

Perceived Acceptability and Feasibility

Almost all participants (> 98%) rated the ARROW Guide, Intervention Plan Form, and Tools/Worksheet to be helpful or very helpful (see Table 3). Approximately one-fifth of participants (22.9%, $n = 38$) endorsed concern that the ARROW could create extra work, whereas the remaining 77.1% did not endorse these concerns or were unsure. Most participants agreed or strongly agreed that the ARROW was easy to use (78.1%, $n = 132$) and efficient (84.7%, $n = 143$).

Study 2: Field Study

Although Study 1 study provided initial support for the ARROW in a controlled setting, it used only a single vignette case, and the results could be impacted by practice effects. To extend this research, we next conducted a field study to examine implementation of the ARROW in real-world practice, this time establishing a higher bar for comparison. More specifically, rather than comparing the ARROW to an entirely unstructured approach to case planning (i.e., a blank space in which to write in a case plan), we compared the ARROW to a short case planning form. In a previous study, we found that this case planning form resulted in significantly better case plans than an entirely unstructured approach (Viljoen et al., 2019). Thus, our goal was to determine if the ARROW could lead to improvements in plans above and beyond an approach that had already achieved some degree of success.

In addition, rather than solely examining formulations and plans, we examined the delivery of interventions in the six months following the plan. This is a novel contribution of the present study. Existing studies on case planning systems have not yet tested whether improvements in plans carry over to the delivery of interventions (e.g., Bosker & Witteman, 2016). However, research suggests that even if professionals develop good plans, these plans are not always carried out (Harris, Gingerich, & Whittaker, 2004; Peterson-Badali et al., 2015).

Method

Participants

Our partners, the provincial youth justice authority, operate 47 youth probation offices serving over 100 communities. As such, our final set of propensity-score matched participants included 320 adolescents who were on probation throughout the province; 160 adolescents were in the ARROW group, and 160 were in the comparison group. The mean age of participants was 17.30 years ($SD = 1.18$ years), and 78.8% of participants were male ($n = 252$). Approximately half were Caucasian (52.7%, $n = 168$), 31.3% ($n = 100$) were Indigenous, 6.6% ($n = 21$) were Asian, 3.8% ($n = 12$) were African or Black, and 1.6% ($n = 5$) were Hispanic. Most participants were charged with a violent offense (44.4%, $n = 142$), property offense (31.9%, $n = 102$), or a

probation violation (31.3%, $n = 100$). Approximately two-thirds of participants had prior convictions (64.4%, $n = 206$), and 42.5% ($n = 136$) had been previously incarcerated.

Procedures

Immediately following the case vignette study and ARROW training, the provincial youth justice authority implemented the ARROW throughout the province (May 2015) by enacting policies wherein YPOs were required to use the ARROW to develop case plans for each adolescent on probation (approximately 1,500 youth/year). This replaced their previous case planning approach (i.e., a simple case planning form). To provide some context for this work, it is important to note that our partner, the youth justice authority, serve a large geographical region (close to 1 million km²), including rural and urban settings, and as such program availability differs by community. However, whereas some programming is locally contracted (e.g., intensive support and supervision programs), all forensic mental health services are delivered through a single provincial agency (with multiple locations); this agency specializes in cognitive-behavioral approaches and provides services to youth who are identified as requiring such services. In addition, all YPOs have training on motivational interviewing.

Comparison sample (i.e., simple case planning form). The pre-ARROW sample, or comparison group, comprised adolescents whose formulations and case plans were completed with the simple case planning form. YPOs were provided with a blank space to write a case formulation. Following this, they selected SAVRY risk and protective factors to include in the plan. The form prompted them to write a goal, planned intervention, and target date for each selected factor.

ARROW sample. The ARROW sample comprised adolescents whose formulations and case plans were completed using the ARROW. Like the simple-case planning form used in the comparison group, the ARROW was completed electronically. However, in contrast to the simple case planning form, this form included instructions for case formulation, as well as suggested intervention strategies from which YPOs could select or add. In addition to the form, all YPOs were provided with the ARROW guide, and nearly all YPOs (93.5%, $n = 102$) attended the ARROW training (described earlier).

Sampling. To sample adolescents in the *comparison* group (i.e., simple case planning form), the youth justice authority generated a list of all adolescents in the province who were on a probation supervision order and had case plans completed in the 8 months following the implementation of the simple case planning form (November 5, 2012 – July 16, 2013). Similarly, to sample participants in the *ARROW* group, the youth justice authority generated a list of all adolescents in the province who were on a probation supervision order and had case plans completed in the 6 months following the implementation of the ARROW (April 30, 2015 – October 15, 2015). We chose these time periods so that the ARROW and comparison group (i.e., simple case planning form) were comparable with respect to the implementation phase; this is important because practices can drift over time. The SAVRY sample was collected over a slightly longer period (8 months vs. 6 months) because the probation offices were less active during the months of November and December. To determine if our study conditions (ARROW and comparison group) differed in any respects other than the ARROW, we reviewed policy documents and consulted with two senior administrators. Based on this investigation, there did

not appear to be any other initiatives (e.g., new training, new evidence-based services, policy changes) that occurred between these time periods that could create confounds.

Coding. From the lists of eligible participants, we randomly sampled participants for the comparison and ARROW groups. Seven trained RAs (five graduate students and two students with a Bachelor's degree) extracted and rated YPOs' formulations, case plans, and records of delivered interventions using a provincial justice database. We excluded cases in which the adolescent did not have a completed case plan (11.8%, $n = 59$). RAs had access to risk assessments, case plans, contact logs, communications with service providers, service referral logs, integrated case management notes, and depending on the client in question, reports from inhouse services (e.g., monthly updates from full-time attendance programs), interim summaries, team updates, court records, breach reports, and presentence reports. Prior to coding, all RAs completed a half-day training on the justice database, which was delivered by a certified trainer. They also received 8 hours of training on the study measures and completed at least five practice cases. To check interrater reliability, a random sample of 62 cases were coded by a second RA with similar training and with access to the same file information.

Propensity score matching. After coding cases, we compared the ARROW sample and comparison group on demographic characteristics (i.e., age, sex, ethnicity), SAVRY total scores, and arrest history (i.e., prior conviction). Adolescents in the ARROW group were significantly older, $t(417) = -3.68, p < .001$, and scored significantly higher on the SAVRY Risk Total score, $t(417) = -3.75, p < .001$. Also, they were significantly more likely to be Indigenous, $\chi^2(N = 419) = 8.69, p = .003$. As such, we conducted propensity score matching in SAS version 9.2 (SAS, 2011) to generate more comparable samples and reduce potential bias (Rosenbaum & Rubin, 1983). Propensity scores were calculated based on age, male sex, ethnicity (i.e., Indigenous or non-Indigenous), prior conviction, probation length, and SAVRY Risk Total score. Six participants were missing data on these items and were thus excluded from the match (Parsons, 2004). Also, to prevent cross-contamination across groups, we eliminated adolescents who were in both the comparison and ARROW groups (2.2%, $n = 11$). The specific matching technique we opted to use was nearest neighbor matching with caliper replacement (Austin, 2011), as this approach appeared to generate equivalent samples. After matching, we did not find any significant differences between samples on covariates (see Table 4).

Missing data. In our final sample of matched cases, there were no missing case plans. However, YPOs did not complete formulations for 41 cases. Formulations were significantly more likely to be completed in the ARROW condition than in the comparison group (i.e., 92.5%, $n = 148$ vs. 81.9%, $n = 131$), $\chi^2(N = 320) = 8.09, p = .004$, perhaps because of the additional guidance that the ARROW provides on how to complete case formulations.

Measures

Adherence Checklist for the Evaluation and Reduction of Risk (ACERR). RAs rated the quality of case formulations and plans using the ACERR (see Study 1). RAs also rated an additional section of the ACERR, the Delivered Interventions section. The items are similar to those in the Plans section of the ACERR but focus on interventions that are delivered in the 6 months following the plan (e.g., whether delivered interventions were the right intensity). ICCs (two-way random effects model, single raters, absolute agreement; McGraw & Wong, 1996)

were in the fair range for Formulations (.53) and Delivered Interventions (.48), and good for Intervention Plans (.62) (Cicchetti, 1994).

Inventory of Interventions. RAs rated the types of interventions in plans using the Inventory of Interventions (see Study 1). ICCs for planned interventions fell in the good to excellent range (i.e., Engagement and Skills = .86, Supervision and Structure = .63, Families and Supports = .75, Services with Research Support = .71, Any Services = .74, Culturally-Relevant Strategies = .65). Using the same set of items, RAs also rated interventions that were delivered in the 6 months following the plan. ICCs for delivered interventions, again, fell in the fair to excellent range (i.e., Engagement and Skills = .66, Supervision and Structure = .47, Families and Supports = .63, Any Services = .77, Culturally-Relevant Strategies = .74). We were unable to examine the extent to which delivered interventions had research support, as the ICCs were inadequate (.26), likely due to lack of information in files.

Data Analysis

To test the impact of the ARROW on formulations, case plans, and intervention-delivery, we conducted analyses of variance (ANOVAs; IBM Corporation, 2013). We interpreted η^2 of .01 as small, .06 as moderate, and .14 as large (Cohen, 1988). Given that some outcome variables showed substantial skew, we ran some of these analyses with a non-parametric alternative, the Mann-Whitney U test, when skew was greater than 1.0. We interpreted an effect size of .10 as small, .30 as moderate, and $> .50$ as large (Cohen, 1988).

Results

Case Formulations

ARROW formulations were rated significantly higher in overall quality (i.e., ACERR Formulation Total) than those in the comparison group (small effects; see Table 5). At an item level, ARROW formulations were more likely to describe protective factors, $F(1, 277) = 18.64$, $p < .001$, $\eta^2 = .06$ (ICC = .76). They were also significantly longer, as measured by the number of sentences, $F(1, 277) = 41.23$, $p < .001$, $\eta^2 = .13$, $U = 5335.00$, $p < .001$ (ICC = .95).

Plans

Compared to plans developed with the simple case planning form, ARROW plans were more likely to include supported and recommended interventions. In particular, they scored higher on Engagement and Skills, Families and Supports, and Services with Research Support; these effect sizes were large (see Table 5). For Indigenous adolescents, ARROW plans scored higher on Culturally-Relevant Interventions (large effects). ARROW plans also scored significantly higher on overall quality (i.e., ACERR Plan Total) and several ACERR items. They were more likely to be tailored to culture, $F(1, 148) = 69.21$, $p < .001$, $\eta^2 = .32$ (ICC = .86), include a broader array of strategies, $F(1, 318) = 41.25$, $p < .001$, $\eta^2 = .12$ (ICC = .54), interventions with research support, $F(1, 318) = 16.53$, $p < .001$, $\eta^2 = .05$ (ICC = .50), and strategies to build or leverage protective factors, $F(1, 318) = 5.11$, $p = .024$, $\eta^2 = .02$ (ICC = .76).

Delivered Interventions

Next, we tested interventions that YPOs recorded as being delivered in the 6 months following the plan. To ensure an equivalent rating period for all adolescents, we restricted these analyses to adolescents who had been on probation for at least 6 months ($n = 292$). In the ARROW sample, adolescents were recorded as receiving more strategies related to Supervision and Structure (small effects; see Table 5). However, few other differences were significant. Although the interrater reliability of specific ACERR items was too low to examine item-level differences (six items had ICCs $< .40$, poor range; Cicchetti, 1994), in the ARROW sample, RAs rated professionals as lower on the extent to which they carried out the strategies listed in their plans, $F(1, 288) = 17.97, p < .001, \eta^2 = .06$ (ICC = .47).

Discussion

Even though risk assessment tools like the SAVRY are intended to guide case planning, they do not provide much guidance. Thus, we developed and evaluated an intervention-planning tool, the ARROW. In the vignette study, we compared ARROW plans to unstructured plans wherein YPOs were provided a blank space to write in their plan. In the field study, we compared ARROW plans to a simple case planning form wherein YPOs were provided with general prompts but not suggested interventions.

Key Findings

Improvements in case plans. Overall, the results suggest that the ARROW helped to improve case plans beyond the SAVRY alone. Consistent with its first objective, the ARROW appeared to increase the inclusion of best practice strategies in case plans (e.g., cognitive-behavioral therapy, motivational interviewing, incentives for positive behavior, time with prosocial supports) (see Landenberger & Lipsey, 2005; Lipsey, 2009; Wong, 2005). Consistent with its second objective, ARROW plans showed a greater range of strategies than non-ARROW plans. Rather than relying on surveillance alone, ARROW plans included strategies that YPOs planned to use in their sessions (e.g., help adolescents problem solve). Consistent with its third objective, ARROW plans included more culturally-tailored strategies for adolescents who are Indigenous. For instance, in the vignette study, only 27.5% of unstructured case plans included culturally-relevant strategies (e.g., consult with cultural experts), compared to 94.5% of ARROW plans. Finally, consistent with its fourth objective, the ARROW was associated with general improvements in case plans; ARROW plans were more detailed, more likely to include protective factors, and more likely to address criminogenic needs.

What is the underlying mechanism for these observed changes in plans? Although this remains to be tested, in our view, these improvements likely occurred because the ARROW decision support system directly integrates recommended practices into the case planning process. For instance, the computerized form populates with suggested interventions. Although YPOs are not required to select these strategies, by listing possible strategies, the ARROW may orient YPOs to interventions that they might not otherwise have considered. In other words, the ARROW appears to serve as a memory and decisional aid. It is also possible that the ARROW reminds YPOs to record strategies that they were already using so that their plans better capture their actual behaviors.

Improvements in case formulation. Although the primary aim of the ARROW is to

improve case plans, results from both studies indicated that the ARROW was also associated with improvements in formulations. For instance, in the vignette study, ARROW formulations were more detailed, more likely to describe how risk and protective factors operate, and more likely to specify the potential types of future offending. That said, many formulations still had limitations even after the ARROW was implemented, indicating that formulation is a difficult task which requires concentrated, continued efforts (e.g., ongoing support and feedback).

Impact on delivered interventions is unclear. Following the implementation of the ARROW, YPOs were slightly more likely to record using strategies to provide supervision and structure, engage the adolescents, and build adolescents' skills. Thus, to some extent, improvements in case plans might carry over to the delivery of interventions. However, overall, we found few significant differences. In addition, in the ARROW condition, RAs rated YPOs as lower on the extent to which they carried out their plans. This could indicate that YPOs' plans were overly ambitious or that evidence-based interventions were not available. Alternatively, this could stem from measurement problems, as we were relying on YPOs' written case logs. For example, although ARROW plans often included strategies that YPOs planned to use in their sessions (e.g., problem solving with youth), YPOs' case logs were brief and YPOs typically did not explicitly describe what interventions strategies they used in their sessions with youth. Thus, even if they were carrying out these strategies, YPOs did not routinely record this in their case logs. Also, even when YPOs referred youth to services, YPOs' files did not consistently provide information on if and how often youth attended services, what specific strategies or approaches service providers utilized (e.g., cognitive behavior approaches or another strategy), and which risk factors service providers targeted (e.g., substance use and/or anger management). Regardless, we believe that consistent with other studies (e.g., Peterson-Badali et al., 2015), these results highlight that good plans do not necessarily guarantee better intervention delivery. Furthermore, we recommend that agencies develop procedures to better record the interventions that they are using, such as standardized templates for case logs, and adopt strategies to facilitate regular communication between YPOs and service providers.

High perceived acceptability. Besides predicting some improvements in plans and formulations, we found that most YPOs and managers viewed the ARROW as useful, and feasible to implement. This is likely because YPOs played an instrumental role in the development of the ARROW. Given that this evaluation occurred immediately following the ARROW training, follow-up research is needed to assess the sustainability of effects.

Limitations

Although both studies have limitations, together they provide a more complete set of results. Our first study, the vignette study, provided a controlled means by which to evaluate the ARROW prior to its adoption. However, similarly to some studies (e.g., Dudley, Park, James, & Dodgson, 2010), it included only a single vignette. Thus, in the field study, we examined a large sample of real-world cases. Although we used propensity-score matching to reduce the possibility of confounds, this approach is not as rigorous as randomized controlled studies. In addition, given that we wished to compare the ARROW and the simple case planning form group immediately after they were implemented (i.e., prior to drift), this meant there was 1.75-year period between the two samples. Although there did not appear to be significant changes in other policies or services during this time, it is nevertheless a potential limitation.

Another limitation is that, although interrater reliability for our study measures was generally good to excellent for the vignette study, it was lower for the field study, falling in the fair to good range, indicating that it may be more difficult to evaluate the quality of plans in real-world contexts. Also, as noted, it was challenging to rate delivered interventions because YPOs did not routinely record strategies that they used in their sessions with youth, nor did we have access to therapists' notes. This suggests a need for better documentation of interventions. Finally, it was not possible to keep raters blind to the study condition, as the ARROW uses an identifiable template for case plans.

Implications for Research and Practice

Although we found some evidence to support the ARROW, researchers must continue to investigate approaches to improve case plans. For instance, although the ARROW is designed to accompany the SAVRY, researchers could develop similar intervention-planning tools to accompany other risk assessment tools. Alternatively, they can design and test other strategies to improve case plans, such as audit systems, or individualized feedback and coaching. Whereas the ARROW provides professionals with considerable discretion to select intervention strategies, researchers can also study the pros and cons of approaches that are more prescriptive.

In evaluating case planning systems, researchers should examine whether professionals view the approach as acceptable and feasible (Proctor et al., 2011). Many initiatives fail to have their desired effect because of implementation problems (Bauer et al., 2015; Vincent, Guy, Perrault, & Gershenson, 2016). As such, acceptability and feasibility are critical. However, approaches that are quick and easy can also have some disadvantages. For instance, on a checklist-style of form, YPOs might quickly check off strategies without careful thought. As such, the ARROW decision support system includes a checklist, but it also prompts YPOs to provide details about how they will implement the selected strategies. However, future research should explore whether YPOs show adequate adherence to this system.

Relatedly, researchers should study the long-term sustainability of case planning approaches. For instance, although we found improvements in the quality of case plans in the six months following the implementation of the ARROW, practices can drift over time (Wiltsey et al., 2012). Researchers should also test whether case planning approaches like the ARROW can be transported to other jurisdictions. For instance, given that we implemented the ARROW in nearly 50 offices, providing YPOs with a single service inventory would have limited transportability, especially as program availability varies by location and changes over time. Instead, the ARROW educates YPOs about which types of interventions have research support. It also provides a broader range of interventions that YPOs can utilize themselves even if evidence-based interventions are not available (e.g., motivational interviewing, arrange for youth to spend time with prosocial supports). However, treatment-oriented case planning systems, like the ARROW, might be challenging to transport into settings in which YPOs view their role as being primarily to monitor probation compliance rather than to administer or deliver treatment.

Although the ARROW includes attention to cultural considerations for Indigenous youth, further research on culture is critical. Researchers should examine how YPOs approach case planning for adolescents who are African or Black, Asian, Indigenous, and Latinx, and/or from sexual, gender, and religious minority groups. To conduct this research, it is important for

researchers to seek guidance and adopt a collaborative approach. For instance, even when it is not their intention, researchers have a history of undermining the expertise and autonomy of Indigenous people and conducting research “on” rather than “with” Indigenous people (Maar et al., 2011). As such, in this project, Indigenous professionals led the development of the ARROW modules pertaining to Indigenous youth and consulted knowledge keepers and Elders with cultural expertise.

Besides examining approaches to improve case planning, researchers need to broaden their focus to examine the other processes involved in risk management (e.g., formulation, intervention delivery, reevaluation of risk), and test how these pieces fit together and influence youths’ outcomes, such as their recidivism rates. For instance, although researchers hypothesize that improved case formulations can improve case planning for offenders or treatment outcomes (Hart et al., 2011), researchers have not yet tested this assertion. Similarly, little research has tested the extent whether improved case plans translate into improvements in intervention delivery. As illustrated in this study and prior research, challenges can arise (Peterson-Badali et al., 2015). For instance, developing a good plan has little value if evidence-based treatment programs are not available. As such, ultimately, researchers need to work toward ensuring that youth have greater access to effective programs. Finally, researchers can explore whether improvements in client outcomes (e.g., recidivism) are mediated or explained by improved needs matching (Luong & Wormith, 2011; Peterson-Badali et al., 2015).

Besides pointing to areas for future research, this project also has several implications for practice. Our results, and the results of other studies (Viljoen et al., 2018), highlight that risk assessment tools alone are insufficient to guarantee better case plans. As such, justice agencies should provide their staff with greater support in case planning, such as by providing them with case planning tools, training, or templates. In addition, justice agencies should be proactive in seeking out ways to evaluate and improve their case planning approaches. For instance, the current project came about because our partners proactively sought our assistance in addressing a real-world challenge – they wanted to ensure that YPOs had the support and guidance necessary to create the best possible case plans. This commitment to improving practices is commendable, and a goal that we hope other agencies can work towards as well.

In sum, even though we now know have a better understanding of what works to reduce adolescent offending, many adolescents on probation do not receive supported interventions. Through this partnership project, we found that intervention-planning tools such as the ARROW may be one strategy by which to improve case plans. However, improving plans is only one step on what is mostly likely a long and complicated pathway between risk assessment and ultimately, the prevention or reduction of reoffending.

References

- Alexander, J. F., Waldron, H. B., Robbins, M. S., & Neeb, A. A. (2013). *Functional Family Therapy for adolescent behavior problems*. Washington, DC: American Psychological Association.
- American Psychological Association. (2006). Evidence-based practice in psychology. *American Psychologist, 61*, 271–285. doi:10.1037/0003-066X.61.4.271
- American Psychological Association. (2017). *Ethical principles of psychologists and code of conduct* (2002, Amended August 3, 2016, Adopted January 1, 2017). Retrieved from <https://www.apa.org/ethics/code/ethics-code-2017.pdf>
- Austin, P. C. (2011). A tutorial and case study in propensity score analysis: An application to estimating the effect of in-hospital smoking cessation counseling on morality. *Multivariate Behavioral Research, 46*, 119-151. doi:10.1080/00273171.2011.540480
- Bauer, M. S., Damschroder, L., Hagedorn, H., Smith, J., & Kilbourne, A. M. (2015). An introduction to implementation science for the non-specialist. *BMC Psychology, 3*, 32. doi:10.1186/s40359-015-0089-9
- Beidas, R. S., Cross, W., & Dorsey, S. (2014). Show me, don't tell me: Behavioral rehearsal as a training and analogue fidelity tool. *Cognitive and Behavioral Practice, 21*, 1-11. doi:10.1016/j.cbpra.2013.04.002
- Bonta, J., & Andrews, D. A. (2017). *The psychology of criminal conduct* (6th ed.). New York, NY: Routledge.
- Bonta, J., Bourgon, G., Rugge, T., Scott, T.L., Yessine, A. K., Gutierrez, L., & Li, J. (2011). An experimental demonstration of training probation officers in evidence-based community supervision. *Criminal Justice & Behavior, 38*, 1127–1148. doi:10.1177/0093854811420678
- Bonta, J., Rugge, T., Scott, T., Bourgon, G., & Yessine, A. (2008). Exploring the black box of community supervision. *Journal of Offender Rehabilitation, 47*, 248-270. doi:10.1080/10509670802134085
- Borum, R., Bartel, P., & Forth, A. (2006). *Manual for the Structured Assessment for Violence Risk in Youth (SAVRY)*. Odessa, FL: Psychological Assessment Resources.
- Bosker, J., & Witteman, C. (2016). Finding the right focus: Improving the link between risk/needs assessment and case management in probation. *Psychology, Public Policy, and Law, 22*, 221-233. doi:10.1037/law0000075
- Bosker, J., Witteman, C., & Hermanns, J. (2013). Do intervention plans meet criteria for effective practice to reduce recidivism? How probation officers forget about social capital and basic needs. *European Journal of Probation, 5*, 65–85. doi:10.1177/206622031300500105

- Bosker, J., Witteman, C., Hermanns, J., & Heij, D. (2015). Improving agreement about intervention plans in probation by decision support. *International Journal of Offender Therapy and Comparative Criminology*, *59*, 1459-1473. doi:10.1177/0306624X14539127
- Bourgon, G., Gutierrez, L., & Aston, J. (2012). The evolution of community supervision practice: The transformation from case manager to change agent. *Federal Probation*, *76*, 27-35.
- Brown, S., & Völlm, B. (2013). Case formulation in personality disordered offenders: Views from the front line. *Criminal Behaviour and Mental Health*, *23*, 263-273. doi:10.1002/cbm.1887
- Canadian Psychological Association (2017). *Canadian code of ethics for psychologists* (4th ed.). Ottawa, ON: Authors.
- Chadwick, N., DeWolf, A., & Serin, R. (2015). Effectively training community supervision officers: A meta-analytic review of the impact on offender outcome. *Criminal Justice and Behavior*, *42*, 977-989. doi:10.1177/0093854815595661
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*, 284-290. doi:10.1037/1040-3590.6.4.284
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20 (Version 3): Assessing risk for violence*. Burnaby, BC: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Douglas, K. S., Webster, C. D., Hart, S. D., Eaves, D., & Ogloff, J. R. P. (Eds.). (2001). *HCR-20: Violence risk management companion guide*. Burnaby, BC: Mental Health, Law, and Policy Institute, Simon Fraser University, and Department of Mental Health Law & Policy, University of South Florida.
- Dudley, R., Park, I., James, I., & Dodgson, G. (2010). Rate of agreement between clinicians on the content of a cognitive formulation of delusional beliefs: The effect of qualifications and experience. *Behavioural and Cognitive Psychotherapy*, *38*, 185-200. doi:10.1017/S1352465809990658
- Dyck, H. L. (2016). *An empirical evaluation of the RCMP's youth intervention and diversion strategy in New Brunswick*. Unpublished doctoral dissertation, University of New Brunswick, Fredericton, New Brunswick, Canada.
- Eccles, M. P., Armstrong, D., Baker, R., Cleary, K., Davies, H., Davies, S., ... & Sibbald, B. (2009). An implementation research agenda. *Implementation Science*, *4*, 18-25. doi:10.1186/1748-5908-4-18

- Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation components. *Research on Social Work Practice, 19*, 531-540. doi:10.1177/1049731509335549
- Fixsen, D. L., Naoom, S. F., Blasé, 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, The National Implementation Research Network. Retrieved from <http://fpg.unc.edu/sites/fpg.unc.edu/files/resources/reports-and-policy-briefs/NIRN-MonographFull-01-2005.pdf>
- Giguère, A., Légaré, F., Grimshaw, J., Turcotte, S., Fiander, M., Grudniewicz, A., & ... Gagnon, M. (2012). Printed educational materials: Effects on professional practice and healthcare outcomes. *The Cochrane Database of Systematic Reviews, 10*, CD004398. doi:10.1002/14651858.CD004398.pub3
- Hannah-Moffat, K., & Maurutto, P. (2003). *Youth risk/need assessment: An overview of issues and practices*. Ottawa, ON: Department of Justice Canada.
- Harris, P. M., Gingerich, R., & Whittaker, T. A. (2004). The “effectiveness” of differential supervision. *Crime & Delinquency, 50*, 235–271. doi:10.1177/0011128703258939
- Hart, S. D., Kropp, P. R., Laws, D. R., Klaver, J., Logan, C., & Watt, K. A. (2003). *The Risk for Sexual Violence Protocol (RSVP): Structured professional guidelines for assessing risk of sexual violence*. Burnaby, BC: Mental Health, Law, and Policy Institute, Simon Fraser University; Pacific Psychological Assessment Corporation; & the British Columbia Institute Against Family Violence.
- Hart, S., Sturmey, P., Logan, C., & McMurrin, M. (2011). Forensic case formulation. *The International Journal of Forensic Mental Health, 10*, 118-126. doi:10.1080/14999013.2011.577137
- Henggeler, S. W., & Schoenwald, S. K. (2011). Evidence-based interventions for juvenile offenders and juvenile justice policies that support them. *Social Policy Report, 25*, 1–28. doi:10.1108/JACPR-01-2015-0155
- Holloway, E. D. (2015). *Risk assessment and case planning by juvenile probation officers: The role of trauma*. (M.A., Fordham University). ProQuest Dissertations and Theses. Retrieved from <http://search.proquest.com.proxy.lib.sfu.ca/docview/1733682593?accountid=13800>
- Hunt, D. L., Haynes, B., Hanna, S. E., & Smith, K. (1998). Effects of computer-based clinical decision support systems on physician performance and patient outcomes. *Journal of the American Medical Association, 280*, 1339-1346. doi:10.1001/jama.280.15.1339
- IBM Corporation (2013). *IBM SPSS statistics for Windows, Version 22.0*. Armonk, NY: IBM Corporation.

- Kirkness, V., & Barnhardt, R. (2001). First Nations and higher education: The four R's – respect, relevance, reciprocity, responsibility. In R. Hayoe and J. Pan (Eds.), *Knowledge across cultures: A contribution to dialogue among civilizations* (pp. 1–17). Hong Kong, CN: Comparative Education Research Centre, University of Hong Kong.
- Landenberger, N. A., & Lipsey, M. W. (2005). The positive effects of cognitive-behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology*, *1*(4), 451-476. doi:10.1007/s11292-005-3541-7
- Lipsey, M. W. (2009). The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic overview. *Victims and Offenders*, *4*, 124-147. doi:10.1080/15564880802612573
- Luong, D., & Wormith, J. (2011). Applying risk/need assessment to probation practice and its impact on the recidivism of young offenders. *Criminal Justice and Behavior*, *38*, 1177-1199. doi:10.1177/0093854811421596
- Maar, M. A., Lightfoot, N. E., Sutherland, M. E., Strasser, R. P., Wilson, K. J., Lidstone-Jones, C. M., ... & Lesage, M. T. (2011). Thinking outside the box: Aboriginal people's suggestions for conducting health studies with Aboriginal communities. *Public Health*, *125*, 747-753. doi:10.1016/j.puhe.2011.08.006
- Mahoney, J. L., & Stattin, H. (2000). Leisure activities and adolescent antisocial behavior: The role of structure and social context. *Journal of Adolescence*, *23*, 113-127. doi:10.1006/jado.2000.0302
- McCart, M. R., & Sheidow, A. J. (2016). Evidence-based psychosocial treatments for adolescents with disruptive behavior. *Journal of Clinical Child and Adolescent Psychology*, *45*, 529-563.
- McCaslin, W. D., & Boyer, Y. (2009). First Nations communities at risk and in crisis: Justice and security. *Journal of Aboriginal Health*, *5*, 61-87. doi:10.3138/ijih.v5i2.28983
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*, 30-46. doi:10.1037/1082-989x.1.1.30
- McMurran, M. (2009). Motivational interviewing with offenders: A systematic review. *Legal and Criminological Psychology*, *14*, 83-100. doi:10.1348/135532508X278326
- Michie, S., van Stralen, M. M., & West, R. (2011). The Behaviour Change Wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, *6*(42), 1-11. doi:10.1186/1748-5908-6-42
- Moja, L., Kwag, K. H., Lytras, T., Bertizzolo, L., Brandt, L., Pecoraro, V., & ... Bonovas, S. (2014). Effectiveness of computerized decision support systems linked to electronic health records: A systematic review and meta-analysis. *American Journal of Public Health*, *104*(12), e12-e22. doi:10.2105/AJPH.2014.302164
- Monchalin, L. (2016). *The colonial problem: An Indigenous perspective on crime and injustice*

- in Canada*. Toronto, ON: University of Toronto Press.
- Nelson, R. J., & Vincent, G. M. (2018). Matching services to criminogenic needs following comprehensive risk assessment implementation in juvenile probation. *Criminal Justice and Behavior, 45*, 1136-1153. doi:10.1177/0093854818780923
- Parsons, L. S. (2004, May). Performing a 1: N case-control match on propensity score. In *Proceedings of the 29th Annual SAS users group international conference* (pp. 165-29). SAS Institute.
- Peterson-Badali, M., Skilling, T., & Haqanee, Z. (2015). Examining implementation of risk assessment in case management for youth in the justice system. *Criminal Justice and Behavior, 42*, 304-320. doi:10.1177/0093854814549595
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., & ... Hensley, M. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research, 38*, 65-76. doi:10.1007/s10488-010-0319-7
- Roberts, J. V., & Reid, A. A. (2017). Aboriginal incarceration in Canada since 1978: Every picture tells the same story. *Canadian Journal of Criminology and Criminal Justice, 59*, 313-345. doi:10.3138/cjccj.2016.E24
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of propensity score in observational studies for causal effects. *Biometrika, 70*, 41-55. doi:10.1093/biomet/70.1.41
- Shepherd, S. M., & Ilalio, T. (2016). Maori and Pacific Islander overrepresentation in the Australian criminal justice system—what are the determinants? *Journal of Offender Rehabilitation, 55*, 113–128. doi:10.1080/10509674.2015.1124959
- Singh, J. P., Desmarais, S. L., Sellers, B. G., Hylton, T., Tirotti, M., & Van Dorn, R. A. (2014). From risk assessment to risk management: Matching interventions to adolescent offenders' strengths and vulnerabilities. *Children and Youth Services Review, 47*, 1-9. doi:10.1016/j.childyouth.2013.09.015
- Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited: An overview of some recommended measures of effect size. *Trends in Sport Sciences, 1*(21), 19-25.
- Trotter, C. (1996). The impact of different supervision practices in community corrections: Cause for optimism. *Australian and New Zealand Journal of Criminology, 29*, 29–46. doi:10.1177/000486589602900103
- Truth and Reconciliation Commission of Canada. (2015). *Honouring the truth, reconciling for the future: Summary of the final report of the Truth and Reconciliation Commission of Canada*. Winnipeg, MB: Truth and Reconciliation Commission of Canada. Retrieved from http://www.trc.ca/websites/trcinstitution/File/2015/Findings/Exec_Summary_2015_05_3

1_web_o.pdf

- 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. doi:10.1177/0093854808331249
- Viljoen, J. L., Brodersen, E., Shaffer, C., & McMahon, R. J. (2016). Risk reduction interventions for adolescent offenders. In K. Heilbrun, D. DeMatteo, and N. S. Goldstein, (Eds.), *APA Handbook of Psychology and Juvenile Justice* (pp. 517-544). Washington, DC, US: American Psychological Association. doi:10.1037/14643-024
- Viljoen, J. L., Brodersen, E., Shaffer, C., Muir, N., & ARROW Advisory Board (2014). *Adolescent Risk Reduction and Resilient Outcomes Work-Plan*. Burnaby, British Columbia: Simon Fraser University.
- Viljoen, J. L., Cochrane, D. M., & Jonnson, M. R. (2018). Do risk assessment tools help to manage and reduce risk of violence and reoffending? A systematic review. *Law and Human Behavior, 42*, 181-214. doi:10.1037/lhb0000280
- Viljoen, J. L., Shaffer, C. S., Muir, N. M., Cochrane, D. M., & Brodersen, E. M. (2019). Improving case plans and interventions for adolescents on probation: The implementation of the SAVRY and a structured case planning form. *Criminal Justice and Behavior, 46*, 42-62. doi:10.1177/0093854818799379
- Vincent, G. M., Guy, L. S., Perrault, R. T., & Gershenson, B. (2016). Risk assessment matters, but only when implemented well: A multisite study in juvenile probation. *Law and Human Behavior, 40*, 683-696. doi:10.1037/lhb0000214.supp
- Wachter, A. (2015). *Statewide risk assessment in juvenile probation: JJGPS StateScan*. Pittsburgh, PA: National Center for Juvenile Justice.
- Wiltsey Stirman, S., Kimberly, J., Cook, N., Calloway, A., Castro, F., & Charns, M. (2012). The sustainability of new programs and innovations: a review of the empirical literature and recommendations for future research. *Implementation Science, 7*(17), 17. doi:10.1186/1748-5908-7-17
- Wong, S. (2005). The effects of adolescent activities on delinquency: A differential involvement approach. *Journal of Youth and Adolescence, 34*, 321-333. doi:10.1007/s10964-005-5755-4

Table 1

Vignette Study: Quality of Case Formulations

ACCERR Total and Items	Unstructured Condition		ARROW Condition		Repeated-Measure ANOVA		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>p</i>	η^2
ACERR Formulation Total	3.41	1.97	7.63	2.52	314.10	<.001	.70
Described risk factors' role in offending	1.17	0.62	1.90	0.64	101.06	<.001	.43
Described protective factors' role ^a	0.04	0.19	1.14	0.87	207.97	<.001	.61
Specified types of offending	0.66	0.68	1.25	0.85	54.14	<.001	.29
Described scenarios	0.60	0.69	1.51	0.75	125.56	<.001	.49
Had adequate detail	0.93	0.63	1.82	0.69	150.61	<.001	.53

Note. $n = 134$ matched formulations. ACCERR = Adherence Checklist for the Evaluation and Reduction of Risk. ^a This item showed skew > 1.0 . As such, we conducted a Wilcoxon Signed-Rank test, which was significant, $Z = -8.69$, $p < .001$, $ES = .52$.

Table 2

Vignette Study: Quality of Case Plans

	Unstructured Condition		ARROW Condition		Repeated-Measure ANOVA			Wilcoxon Signed- Rank Test		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>p</i>	η^2	<i>Z</i>	<i>p</i>	<i>ES</i>
ACERR Plan Total and Items	9.64	3.46	14.30	2.71	196.39	<.001	.57	–	–	–
Targeted key risk factors	1.44	0.62	1.89	0.62	43.12	<.001	.23	–	–	–
Was appropriate intensity	1.46	0.63	1.58	0.58	3.15	.078	.02	–	–	–
Included best practices	1.29	0.62	2.05	0.70	114.65	<.001	.44	–	–	–
Included protective factors	0.50	0.70	1.45	.89	105.69	<.001	.42	–	–	–
Tailored plan to culture ^a	0.37	0.67	1.97	0.48	–	–	–	-10.36	<.001	.75
Included array of strategies	0.95	0.69	1.90	0.54	194.05	<.001	.57	–	–	–
Had manageable focus ^a	1.35	0.63	1.72	0.55	–	–	–	-4.81	<.001	.19
Had adequate detail	0.98	0.71	1.61	0.60	75.25	<.001	.34	–	–	–
Inventory of Interventions										
1. Engagement and Skills ^a	0.48	1.38	6.71	2.42	–	–	–	-10.44	<.001	.96
2. Supervision and Structure ^a	0.82	1.21	2.84	1.99	–	–	–	-8.43	<.001	.53
3. Families and Supports ^a	1.26	1.64	5.51	2.96	–	–	–	-9.38	<.001	.72
4. Services with Research Support ^a	0.34	0.78	3.28	1.75	–	–	–	-10.19	<.001	.81
5. Any Services	6.10	2.43	5.17	2.73	11.77	.001	.08	–	–	–
6. Culturally-Tailored Strategies	0.35	0.87	3.15	2.31	–	–	–	-9.50	<.001	.66

Note. $n = 147$ matched plans. ACERR = Adherence Checklist for the Evaluation and Reduction of Risk; Z = Z -coefficient from Wilcoxon Signed-Rank test (Z -coefficients are always negative); ES = effect size of Wilcoxon Signed-Rank test. * $p < .05$, * $p < .01$, *** $p < .001$. ^a Skew > 1.0 ; in these cases, the Wilcoxon Signed-Rank Test provides a more appropriate test of associations than the repeated-measure ANOVA.

Table 3

Vignette Study: Perceived Acceptability of the ARROW

	Very Helpful		Helpful		Not That Helpful		Very Unhelpful	
	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>
Will ARROW be useful to YPOs?	36.7	61	62.0	103	1.2	2	0	0
Will ARROW be useful to you?	31.5	52	63.0	104	5.5	9	0	0
How helpful are:								
ARRROW Intervention Plan	34.3	57	65.7	109	0	0	0	0
ARROW Guide	43.1	72	56.3	94	0.6	1	0	0
ARROW Worksheets/Tools	42.5	71	55.7	93	1.8	3	0	0

Note. YPOs = youth probation officers.

Table 4

Sample Characteristics Before and After Propensity Score Matching: Nearest Neighbor Caliper Matching

	Before Match		After Match	
	SAVRY/Plan (n = 167)	ARROW (n = 252)	SAVRY/Plan (n = 160)	ARROW (n = 160)
Mean age	17.12 [†]	17.56 [†]	17.20	17.40
% Male	74.9	81.7	75.0	82.5
% Indigenous	27.5 [†]	41.7 [†]	28.8	33.8
% Prior conviction	68.9	61.1	68.8	60.0
Mean probation length (days)	379.08	379.52	374.04	381.08
Mean SAVRY Risk Total Score	19.40 [†]	23.07 [†]	19.95	21.36

Note. SAVRY = Structured Assessment of Violence Risk in Youth. [†] Signifies a significant difference between SAVRY/Plan and ARROW samples.

Table 5

Field Study: Case Formulations, Plans, and Delivered Interventions

Outcome Measures	Comparison Group (<i>n</i> = 160)		ARROW Group (<i>n</i> = 160)		ANOVA			Mann-Whitney U Test		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>p</i>	η^2	<i>U</i>	<i>p</i>	<i>r</i>
Formulations (<i>n</i> = 279)										
ACERR Formulation Total	5.69	2.33	6.35	2.29	5.74	.017	.02	–	–	–
Plans (<i>n</i> = 320)										
ACERR Plans Total	13.52	5.59	15.18	5.59	7.09	.008	.02	–	–	–
Inventory of Interventions										
1. Engagement & Skills ^a	0.74	1.26	4.24	2.99	–	–	–	4090.00	<.001	.62
2. Supervision & Structure ^a	1.47	1.47	2.00	1.93	–	–	–	11017.00	.032	.12
3. Families & Supports ^a	1.44	1.68	3.19	2.69	–	–	–	7444.50	<.001	.36
4. Any Services	3.48	2.60	3.90	2.89	1.91	.167	.01	–	–	–
5. Services with Research Support ^a	0.25	0.63	1.51	1.34	–	–	–	5903.50	<.001	.54
6. Culturally-Tailored Strategies ^{a, b}	0.43	0.90	2.85	2.21	–	–	–	302.00	<.001	.62
Delivered Interventions (<i>n</i> = 292)										
ACERR Interventions Total	13.19	4.58	12.07	5.21	3.84	.051	.01	–	–	–
Inventory of Interventions										
1. Engagement & Skills ^a	1.10	1.36	1.58	1.92	–	–	–	9415.00	.069	.11
2. Supervision & Structure	3.58	1.95	4.17	2.10	6.22	.013	.02	–	–	–
3. Families & Supports ^a	1.56	1.62	1.64	1.99	–	–	–	10261.50	.574	.03
4. Any Services	4.19	2.97	4.73	3.34	2.07	.151	.01	–	–	–
5. Culturally-Tailored Strategies ^{a, b}	0.54	0.90	1.15	2.23	–	–	–	646.50	.509	.08

Note: ACERR = Adherence Checklist for the Evaluation and Reduction of Risk; U = Mann Whitney U test. ^a Skew > 1.0; in these cases, the Mann-Whitney U test is a more appropriate test than the ANOVA. ^b We restricted analyses on culturally-tailored strategies to adolescents who are Indigenous (*n* = 85 for plans and 75 for delivered). We were unable to examine Services with Research Support for *Delivered* Interventions as ICCs were inadequate.