

**INTEGRATING SCIENCE AND POLICING: EVALUATING AND
EXPLORING INTIMATE PARTNER VIOLENCE RISK ASSESSMENT IN
FRONTLINE POLICE WORK IN CANADA**

by

Erin J. deJong

Bachelor of Arts (Honours), Carleton University, 2013

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of

Doctor of Philosophy

in the Graduate Academic Unit of Psychology

Supervisor: Mary Ann Campbell, Ph.D., Department of Psychology

Examining Board: Nicole O’Byrne, Ph.D., Faculty of Law
Janine Olthius, Ph.D., Department of Psychology
Scott Ronis, Ph.D., Department of Psychology

External Examiner: Mark Olver, Ph.D., Department of Psychology, University of
Saskatchewan

This dissertation is accepted by the
Dean of Graduate Studies

THE UNIVERSITY OF NEW BRUNSWICK

April 2022

© Erin Jacoba deJong, 2022

ABSTRACT

Previous research has found that when police use risk appraisal information to match higher-risk cases of intimate partner violence (IPV) with higher-intensity risk management, it can lead to a decreased likelihood of an individual perpetrating such incidents in the future (Storey et al., 2014). For this reason, many police services across Canada have adopted evidence-based risk instruments that have been shown to predict a perpetrator's risk for future IPV. The Ontario Domestic Assault Risk Assessment (ODARA) is one such tool that is commonly used by Canadian police services because it was specifically designed to be easily scored using information readily available to police officers in the course of their routine investigations (Hilton et al., 2004). However, the ODARA has not yet been evaluated in a frontline context. In addition, its emphasis on static risk (as opposed to dynamic risk) can limit its utility. Using a sample of 309 archival IPV police incident reports drawn from a New Brunswick police service, the current study sought to (a) evaluate police field use of the ODARA over a 12-month follow-up period, and (b) develop and evaluate the validity of a dynamic supplement to the ODARA that could be used by officers to better detect fluctuations in risk and inform risk management tactics. The evaluation results demonstrated problems with officer scoring of the ODARA, which substantially limited the utility of the tool for informing risk management and mitigation. The analyses of dynamic risk factors only supported one new risk item (capturing perpetrator antisocial attitudes), thus preventing the development of a dynamic supplement. The implications for enhancing police use of the ODARA in the field and methodological lessons to consider in future research are discussed.

Keywords: intimate partner violence; risk assessment; recidivism; Ontario Domestic Assault Risk Assessment (ODARA); dynamic risk factors; evidence-based policing

DEDICATION

This dissertation is dedicated to my Opa, Wieger deJong. Thank you for always asking the hard questions, instilling in me a love of learning, encouraging my studies, and reminding me of what is most important in life. Ere Zij God.

ACKNOWLEDGMENTS

This dissertation process has been unlike anything I have undertaken before. Although it has often felt like an uphill marathon, there are many people who have helped me trudge over the finish line through their mentorship, advice, and practical help. Of these people, I first would like to express my deepest appreciation to Jessica McTague, my research partner. Without your sacrifices, generosity, perseverant hard work, patience, and friendship, this dissertation truly would not have been possible. I would also like to extend my sincere appreciation to Angela Totten, our consultant and collaborator, whose contributions to this project have been immeasurable. To Adam and Christa, I thank you from the bottom of my heart for your insightful questions, desire to learn, and dedication to the project. I am also grateful to the many friends and family who have been cheerleaders along the side lines. Your encouragement, curiosity, prayers, and support have helped sustain and motivate me when I have felt too tired to continue. I am grateful for the feedback and advice of my committee members, and most especially my supervisor, Dr. Mary Ann Campbell. Mary Ann, your scholarly expertise and professional contributions have been essential to this dissertation, but also your personal support has meant so much to me throughout this process. I am so grateful for the flexibility, patience, motivation, mentorship, advocacy, teaching, network sharing, generosity, and kindness you have shown me over these past years of graduate school.

Table of Contents

ABSTRACT	ii
DEDICATION.....	iv
ACKNOWLEDGMENTS	v
List of Tables	ix
List of Figures.....	xi
List of Appendices.....	xii
Introduction.....	1
The History of Police Response to IPV in Canada.....	6
The Evolution of General Risk Assessment	26
Dynamic Risk Factors.....	33
IPV Risk Assessment.....	48
Police Use of IPV Risk Assessment Instruments	56
Current Study.....	59
Method	65
Sample	65
Measures	71
Interrater Reliability.....	78
Procedure	80

Results.....	81
Case Descriptive Statistics.....	81
IPV Offender Characteristics.....	82
Victim Characteristics.....	84
IPV Offense Characteristics.....	85
Evaluation of Police Use of the ODARA	86
Development and Evaluation of the Dynamic Supplement to the ODARA.....	102
Discussion.....	114
The Role of Gender.....	115
Police Use of the ODARA.....	122
Development and Evaluation of a Dynamic Supplement to the ODARA.....	138
Evaluation of the ODARA + AA.....	149
Implications and Conclusions.....	151
Implications for Police Use of the ODARA	151
Implications Regarding Dynamic Risk Factors for IPV.....	154
Limitations	156
An Addendum: Considerations about Race and Future Research.....	158
References.....	162
Appendix A.....	243

Appendix B.....	247
Appendix C.....	256
Appendix D.....	259
CURRICULUM VITAE	

List of Tables

Table 1. Results of Chi-square Test and Descriptive Statistics for Perpetrator Characteristics, Overall and as a Function of Perpetrator Gender	222
Table 2. Results of Chi-square Test and Descriptive Statistics for Victim Characteristics, as a Function of Victim Gender.....	223
Table 3. Results of Chi-square Test and Descriptive Statistics for Type of Abuse Committed During the Index Offense by Gender	224
Table 4. Descriptive Statistics and Results of ANOVAs Comparing Males and Female Suspects on Number of Different IPV Actions Perpetrated Within Each Abuse Category	225
Table 5. Results of Chi-square Test and Descriptive Statistics for Reason for the Altercation During the Index Offense by Gender	226
Table 6. Frequency and Percent of IPV Files with Police-Scored ODARA at the Index and Subsequent Recidivism within 12-Month Follow Up Period.....	227
Table 7. Interrater Reliability Estimates between Police Officers and Researchers for ODARA Total Scale Scores (ICC) and Item Scores (Kappa).....	228
Table 8. Summary of Logistic Regression Analysis for Suspect and Contextual Details in Predicting Police Use of ODARA at Index (N = 191).....	229
Table 9. Predictive Validity of ODARA Total Score for Predicting IPV Recidivism as Measured by ROC Analyses	230
Table 10. Number of Actions Taken by Police as a Function of Suspect Gender: Results of ANOVAs and Descriptive Statistics.....	231

Table 11. Results of Chi-square Test for Risk Management Strategies by Police-Rated Risk Classification (Post 2015 files with Police-Scored ODARAs).....	232
Table 12. Kendall’s Tau-b Correlations Between Dynamic Risk Factors and Dichotomous IPV Recidivism Outcome	233
Table 13. Mean Survival Time for Three Levels of Dynamic Risk Factors	234
Table 14. Reliable Change Index Summary Statistics for Dynamic Risk Factors	235
Table 15. Predictive Validity of ‘Antisocial Attitudes’ Dynamic Risk Item for Predicting IPV Recidivism as Measured by ROC Analyses, Compared by Suspect Gender.....	236
Table 16. Predictive Validity of ‘Antisocial Attitudes’ Dynamic Risk Item for Predicting IPV Recidivism as Measured by ROC Analyses, Compared to Researcher-Completed ODARA Total Score.....	237
Table 17. ROC Analysis of Incremental Validity of Adding the ‘Antisocial Attitudes’ Dynamic Risk Item to the ODARA for Predicting IPV Recidivism.....	238
Table 18. Direction of Change in ODARA Risk Score from Index to First Recidivism Event as a Function of Risk Classification at Index.....	239
Table 19. Direction of Change in ODARA + AA Score from Index to First Recidivism Event as a Function of Risk Classification at Index.....	240

List of Figures

Figure 1. ROC Curve for Police-Rated ODARA Score and Recidivism	241
Figure 2. ROC Curve for Researcher-Rated ODARA Score and Recidivism.....	242

List of Appendices

Appendix A. ODARA Item Scoring and Types of Risk Factors.....	243
Appendix B. Coding Guide for IPV Incidents.....	244
Appendix C. Results of Planned Analyses for the Dynamic Risk Composite	256
Appendix D. Predictive Validity of the Dynamic Risk Composite.....	259

Integrating Science and Policing: Evaluating and Exploring Intimate Partner Violence Risk Assessment in Frontline Police Work in Canada

Introduction

Police officers have many roles and responsibilities, and they need to make high-stakes decisions every day in the course of their duties. One scenario that is familiar to many police officers is responding to calls of intimate partner violence (IPV), which makes up approximately 30% of police-reported violent crime in Canada (Burczycka, 2018; Conroy, 2021). IPV is considered actual or threatened abuse by a current or former spouse, common-law, or dating partner, and it can include physical abuse, sexual abuse, psychological/emotional abuse, stalking, and threats of violence (Dawson & Hotton, 2014; Whitaker & Lutzker, 2009). Police officers often experience what is referred to as the “revolving door of domestic violence,” facing the reality that IPV can be a repetitive cycle between the individuals involved (Horwitz et al., 2011). Rates of reoffending for IPV range across studies between 15% (with a 12-month follow-up period) and 60% (with a 10-year follow-up period; Hilton, Harris, Popham, & Lang, 2010; Hilton, Harris, & Rice, 2010; Klein & Tobin, 2008; Loinaz, 2014; Petersson & Strand, 2017). Approximately half of people who have experienced IPV report being abused once in the past five years, one-third report being victimized between two and 10 times in the past five years, and one-fifth report being victimized more than 10 times in that same time frame (Statistics Canada, 2016).

Overall, as many as 44% of Canadian women and 36% of Canadian men who have ever been in an intimate relationship report having experienced IPV at least once in their lifetime (i.e., since the age of 15; Cotter, 2021). However, most cases of IPV are

not reported to the police, and there is a significant gender discrepancy amongst those for whom the police are called (Conroy, 2021). That is, 79% of incidents of IPV that are reported to the police involve a male suspect¹ and female complainant (Department of Justice, 2018a). The Canada-wide rates of police-reported IPV have increased for five consecutive years between 2014 and 2019, up by 14% overall (Conroy, 2021). In addition, 17 police services across Canada have released more recent data showing an average increase of 12% in police-reported IPV from between March and June of 2019 compared to the same four months in 2020 (with speculation that COVID-19 pandemic-related lockdowns have had a large contribution to such increases; Statistics Canada, 2020a). Clearly the problem of IPV is as relevant now as it has ever been, and even small reductions in the rates of IPV can have a meaningful impact on the lives of thousands of women and men every year.

The negative effects of IPV are well documented, with its consequences affecting the social, physical, and psychological wellbeing of victims and their families, friends, and employers (Du Mont et al., 2005). Social consequences of IPV can include the undermining or restricting of opportunities that can result in the victimized person's financial and residential dependence on the perpetrator (Miller & Du Mont, 2000). Physical consequences are exhibited in a vast array of possible health problems, such as cuts, bruises, broken bones, head/spinal cord injuries, internal bleeding, miscarriages, sexually transmitted infections, and potential death (Dal Grande et al., 2003; Dunkle et al., 2004; Krug et al., 2002; Tjaden & Thoennes, 2000). Psychological impacts of IPV

¹ For the sake of brevity and consistency with previous literature, people who were investigated for complaints of IPV will be referred to as *perpetrators* or *suspects*, and those who were their current or former intimate partners as *victims* or *complainants* throughout this dissertation.

include increased victim experiences of depression, anxiety disorders, problematic substance use, eating disorders, posttraumatic stress disorder, suicidal ideation, and a variety of negative emotions (e.g., shame, guilt, fear, low self-esteem; Coker, Davis, et al., 2002; Coker, Smith, et al., 2002; Koss, 1990; Krug et al., 2002; Rodgers, 1994). There are also financial consequences for individuals, communities, and Canadian society in general. This includes expenses associated with emergency shelters, medical expenses, work loss, transition homes, police services, correctional services, social services, income assistance, moving costs, childcare and pet boarding, and mental health services (Department of Justice, 2017a). Previous research has examined the costs of IPV to Canadians, with estimates ranging from \$385 million per year to \$7.4 billion per year (Day, 1995; Greaves et al., 1995; Kerr & McLean, 1996; Varcoe et al., 2011; Zhang et al., 2012).

Police play a crucial role in addressing IPV, often acting as gatekeepers to the criminal justice system and various other social services for both the perpetrators and victims (International Association of Chiefs of Police, 2015; Schonrock & Fawbush, 2018; Stewart, 2001). Police are also in a unique position of being able to influence the judicial outcomes for perpetrators (e.g., offer advice, issue a warning, arrest, lay or recommend charges, support bail, and so on). In fact, police in Canada are actively involved in risk management, and they are becoming increasingly engaged in intervention planning and violence prevention (Public Safety Canada, 2015; Royal Canadian Mounted Police [RCMP], 2019). Previous research has demonstrated that when police use evidence-based principles to guide their decision-making, it can have a positive impact on those involved and lead to a decreased likelihood of the perpetrator

reoffending (Andrews & Bonta, 2003; Andrews et al., 1990; Belfrage et al., 2011; Bonta & Andrews, 2007; Storey et al., 2014; Ward et al., 2007). In particular, research has shown that criminal justice intervention and case management that follows the “Risk-Need-Responsivity” model of offender rehabilitation (a.k.a., the “RNR model”) leads to decreased levels of recidivism in adult male offenders, adult female offenders, and young offenders in both community and custodial settings (Andrews & Bonta, 2010).

The RNR model incorporates cognitive social learning theory and provides theoretical principles to enhance interventions with offenders, describing three core principles (Bonta & Andrews, 2007). First, the risk principle suggests that the intensity of the intervention an offender receives must be matched to their likelihood for recidivism. Mismatching elevates recidivism risk, or at best has neutral effects (Bonta & Andrews, 2017). Second, the need principle is based on the premise that certain risk factors are amenable to change (i.e., dynamic risk factors, or criminogenic needs) and should be targeted in intervention in order to reduce risk for recidivism (e.g., family/marital problems, employment/education problems, problematic substance use, procriminal attitudes; Andrews & Bonta, 2003). Finally, the responsivity principle suggests the intervention the offender receives must be evidence-based (general responsivity) and must be tailored to the learning style, abilities, motivation, and strengths of the offender (specific responsivity; Andrews & Bonta, 2003; Bonta & Andrews, 2017).

Applying the RNR model to policing IPV involves frontline officers using empirically-supported methods to estimate the offenders’ likelihood of reoffending. Police officers should then use this information to guide the intensity of their direct

intervention (e.g., arrest, hold, issue conditions, recommend charges) and to inform recommendations they offer to the perpetrator, victim, Crown attorney, child welfare services, victim services, judge, probation officers, or any other service provider who reads their report or hears their testimony. A key process involved in the application of RNR model is risk assessment, which is a decision-making process that helps determine the best course of action by identifying relevant risk factors to estimate risk of future criminal behaviour (Bloom et al., 2005; Northcott, 2012).

Many risk-assessment measures, some designed specifically for police use, have been developed to structure and improve the accuracy of risk estimates. The use of risk assessment measures has been increasingly embraced by police forces across Canada, both generally and specifically for IPV-related cases (Northcott, 2012). Although risk assessment measures originally tended to focus solely on providing accurate risk estimates, more modern measures also identify criminogenic needs and responsivity factors, which is referred to as dynamic risk assessment. However, there is a need for these more advanced, dynamic risk tools to be developed in a police-accessible format (i.e., quick, easy to fill out, free from technical jargon, and only requiring information police can access in the course of their regular duties; Hilton, Harris, & Rice, 2010). In addition, there is a need for research to examine how police officers are using risk-assessment measures to guide their decision-making and whether offender outcomes improve as a result (i.e., decreasing rates of reoffending). The current dissertation sought to address these needs, starting with a review of the history of police response to IPV in Canada to understand the current context, pressures, and expectations of police officers when responding to calls of IPV.

The History of Police Response to IPV in Canada

Pre-1980s

IPV was historically viewed as a private family matter rather than a public concern (Jaffe et al., 1986). Prior to the 1960s, IPV was grossly overlooked by the criminal justice system (Ogle & Jacobs, 2002). In the 1960s and 1970s, the official position of the criminal justice system was to discourage arrests, charges, and prosecutions for IPV (Garner & Maxwell, 2009). For example, Canadian criminal law was written in such a way that a husband could not be charged with raping his wife (and male victims or those in same-sex relationships were not recognized as potential victims of IPV at all; Martin, 1998; Pattavina et al., 2007; Whitehead et al., 2021). Officers responding to IPV were encouraged to use informal methods of responding, such as mediation and verbal warnings, to resolve “conflict such as that which occurs so frequently between husband and wife” (American Bar Association, 1980, p. 176). Police were trained to only arrest a suspect if (a) they were caught in the middle of hitting the victim, or (b) the victim sustained injuries serious enough to warrant “a certain number of stitches” (Department of Justice, 2002, p. 9).

The 1980s

IPV discourse and policy began to progress in the 1980s, influenced by the second wave of the feminist movement and a surge of research on police response to IPV (Jaffe et al., 1986; Leisenring, 2008; Martin, 1998; Schneider, 2000). The feminist movement of the time contributed to a substantial increase in awareness and interest in how the criminal justice system responds to IPV through grassroots women’s movements and feminist activism (Department of Justice, 2015; Martin, 1998). The

“battered women’s movement” directly challenged the mainstream view that IPV was a private matter between a husband and a wife, instead portraying it as a social and political problem with roots in sexism and oppression (Leisenring, 2008; Schechter, 1982). The efforts of this movement resulted in many shelters, safe houses, crisis lines, and support groups being established for female victims of IPV, and in 1983 Bill C-127 came into effect making sexual assault against one’s wife an offence (*Criminal Code*, 1985, c C-46, s 271(a); Leisenring, 2008). Feminist scholars and activists implored the government to stop overlooking violence against women, yet they continued to overlook the experiences and needs of male victims and those in same-sex relationships (Schneider, 2000). In turn, small changes began to take place in the approach to the policing of IPV that positioned the police to have a more active role in addressing IPV. For example, in 1981, the London Police Department in Ontario became the first Canadian police service to implement policy mandating that police lay charges in all cases of spousal assault (Department of Justice, 2015).

In 1982, the House of Commons set forth a motion that encouraged all Canadian police services to adopt a practice of laying charges for IPV as they would for any other case of common assault (Department of Justice, 2015). One month later, the Solicitor General of Canada wrote a letter to the Canadian Association of Chiefs of Police issuing a directive to lay charges in all cases of IPV in which the “facts and circumstances warrant this action” (Department of Justice, 2015; Jaffe et al., 1986, p. 38). By 1985, policy that promoted formal police intervention in cases of IPV was implemented in nearly every jurisdiction in Canada (Dawson & Hotton, 2014). Many jurisdictions decided to implement “pro-arrest” policies, which strongly encouraged arrests in most

cases while also ultimately offering police some discretion in their decision (Leisenring, 2008; Schneider, 2000). Alternatively, some jurisdictions implemented “mandatory arrest” or “no-drop” policies (Schneider, 2000). Mandatory arrest policies require the arrest of suspects in all cases where there are reasonable grounds to indicate IPV occurred (although “reasonable grounds” was still decided upon by the discretion of the officer). No-drop policies, on the other hand, mandate both arrest and prosecution of suspects for which reasonable grounds can be established (Leisenring, 2008).

Soon after these changes in policy took place in Canada during the 1980s, researchers began to examine their impact. For example, researchers from London, Ontario examined the impact of the original 1981 policy change that mandated laying charges for all IPV cases within that jurisdiction, finding there was a significant increase in police-laid charges (Jaffe et al., 1986; Jaffe et al., 1993). Specifically, they found that charges increased from 3% of reported IPV cases in 1979 to 67% in 1983 and increased further to 89% in 1989 (Jaffe et al., 1986; Jaffe et al., 1993). IPV victims in their studies reported having increased satisfaction with police response and decreased experience of intimate partner abuse after the mandatory charge policies were introduced in 1981 (Jaffe et al., 1986). They also found that victims were significantly more likely to report facing new threats/assaults from their intimate partners when no charges were laid (compared to when charges were laid; Jaffe et al., 1986).

Although the same widespread policy changes had not yet taken place in the United States in the early 1980s, important research was taking place there that would eventually come to have a substantial influence on policy for how the Canadian criminal justice system responds to IPV. In particular, Sherman and Berk (1984) examined how

differential police responses to IPV influenced rates of IPV recidivism in Minneapolis. Beginning in 1981, whenever an officer from the Minneapolis Police Department was called to an IPV incident, the police response was randomly assigned to one of three possible outcomes: (1) arrest the suspect, (2) provide advice (informal mediation), or (3) issue an order for the suspect to leave the home for eight hours. A total of 330 IPV cases were included in the Minneapolis Domestic Violence Experiment. Victim interviews and police records were used in a six-month follow-up period to measure reoffending. They found that rates of reoffending were lower among arrested suspects compared to rates of those advised or removed from the home (Sherman & Berk, 1984). Based on these findings, Sherman and Berk (1984) concluded that “the arrest intervention certainly did not make things worse and may well have made things better” (p. 269). Policy makers interpreted these findings to mean “arrest is best,” which contributed to widespread changes in policies regarding criminal justice system response to IPV in both the United States and Canada (Gartin, 1995).

The 1990s

Despite the fact that there was widespread implementation of pro-arrest, mandatory arrest, and no-drop policies in the 1980s, they were not without controversy. The 1990s saw a surge of debate, literature, and research on the value of policies that mandate a uniform criminal justice system response to IPV. Proponents of these policies argued that pro- and mandatory-arrest policies promote both general and specific deterrence: general deterrence by sending a message to the community that IPV is a serious public issue, and specific deterrence by seeking to prevent IPV reoffending for individual perpetrators (Department of Justice, 2015; Schneider, 2000). The early

research in London, Ontario and Minneapolis, Minnesota was often cited in support of specific deterrence, as it seemed to indicate these policies increased victim satisfaction with police response and decreased IPV reoffending (Jaffe et al., 1986; Jaffe et al., 1993; Sherman & Berk, 1984).

Many feminist scholars of the time argued in support of these policies as a means of holding the police and the government responsible for protecting its citizens (Stark, 2004). Some also argued that having policies that mandate arrest or charging of IPV perpetrators takes the pressure off victims to make the decisions related to police intervention, which they may choose to avoid for various reasons (Stewart, 2001). For example, the implementation of these policies decreases the victim's perception that the perpetrator will hold them responsible for their arrest and thus retaliate in the future. Finally, it was argued that these policies would minimize the subjectivity in police decision-making, which may be inappropriately affected by an officer's personal biases or beliefs about IPV (Avakame & Fyfe, 2001).

In contrast to those who argued in favour of pro-arrest, mandatory arrest, and no-drop policies, many critics and researchers made strong cases against such policies. Criticisms were spurred by the fact that very few studies had successfully replicated the early research findings that supported these policies, despite multiple attempts. Many studies instead showed mixed results, with reoffending rates varying as a function of the length of follow-up time (e.g., within a month of arrest versus more than a month), source of data (e.g., victim versus police report), and perpetrator characteristics (e.g., race, employment status; Belknap, 2001; Berk et al., 1992; Dunford et al., 1990; Hirschel et al., 1992; Schmidt & Sherman, 1996; Sherman, 1992; Sherman, Smith, et al.,

1992). A few studies even found that policies mandating police intervention can increase reoffending by some perpetrators (Dugan et al., 2003; Schmidt & Sherman, 1996; Sherman, Schmidt et al., 1992). In fact, Sherman (1992) himself reviewed attempted replications of the Minneapolis Domestic Violence Experiment and concluded that mandatory arrest laws should be repealed, citing that arrest: (a) only reduces IPV recidivism in some cities and not others, (b) only decreases IPV for certain groups of people and not others, and (c) may reduce short-term IPV risk, but may increase long-term risk.

Feminist critics of pro-arrest, mandatory arrest, and no-drop policies have argued that taking the decision-making completely away from victims of IPV is not always in their best interest (Dasgupta, 2003; Hilton, 1993; Mills, 2003). They argued that blanket policies for criminal justice system response to IPV were not effective at meeting the diverse needs of IPV victims, and they might prevent victims from acting in their own best interest (Crenshaw, 1994; Epstein, 1999). For example, someone who has experienced IPV may not want the perpetrator to be arrested because they might be dependent on them financially or for housing, fear retaliation or escalation of violence, distrust the criminal justice system (especially among visible minorities who may have had negative or unfair experiences with the criminal justice system in the past), fear that they or their partner will be deported (for both legal and illegal immigrants), prefer support and counselling instead of punishment, or not want the relationship to end (Crenshaw, 1994; Epstein, 1999; Leisenring, 2008; Sparks, 1997).

Research on victim perception of mandatory policies supported the notion that a one-size-fits-all response to IPV had the potential to disempower victims and may not

serve everyone equally. Smith (2001) surveyed emergency shelters for female victims of IPV and found that support for mandatory arrest policies varied by race, marital status, and the level of injury incurred. They found that Black women were significantly less likely than White women to support mandatory arrest policies. Similarly, single or divorced women were far more likely to support these policies than those women who were still in a relationship with the perpetrator. Similarly, other research found that women who are financially dependent on their partners are far less supportive of mandatory arrest laws (Leisenring, 2008). One victim expressed: “I need him to pay the bills more than I need him to pay for the violence” (Bohmer et al., 2002, p. 82). Thus, critics argued that IPV victims are too heterogeneous a group for any one uniform police response to be appropriate; instead, they suggested that police response should be tailored to meet the unique needs of the victims of IPV in any given situation (Crenshaw, 1994; Epstein, 1999).

In parallel with the argument that the victims of IPV are too diverse for blanket policies to be appropriate, researchers demonstrated that IPV perpetrators are also a highly heterogeneous group that may require tailored criminal justice system intervention (Stewart, 2001). For example, Sherman et al. (1992) found that “stake in conformity” mediates the effectiveness of arrest in reducing IPV recidivism. That is, arrest resulted in *increased* IPV recidivism for those with a low stake in conforming to societal norms (i.e., the unemployed and unmarried men in their sample). Findings like this called into question whether different identifiable types of IPV perpetrators exist and what implications this might have for policing of IPV. To this effect, Holtzworth-Munroe and Stuart (1994) were among the first researchers to comprehensively review

the existing literature on IPV perpetrators and identify three primary subtypes of IPV perpetrators: relationship/family-only, dysphoric/borderline, and generally violent/antisocial perpetrators. This typology has since become one of the most influential and referenced typologies of IPV perpetrators in the field (Langhinrichsen-Rohling et al., 2000).

Holtzworth-Munroe and Stuart (1994) described the relationship/family-only perpetrators as the least severe offender group, with less severe physical violence (i.e., less likely to result in victim injury) that is limited to family members and with no underlying psychopathology (making up about 50% of IPV perpetrators). The dysphoric/borderline IPV perpetrators engage in moderate to severe IPV, including psychological and sexual abuse (making up about 25% of IPV perpetrators). Their violence is mostly confined to the family, but they may demonstrate other criminal behaviour as well. In addition, these perpetrators tend to be emotionally volatile and may have personality or substance use disorders (Holtzworth-Munroe & Stuart, 1994). The third cluster of IPV perpetrators, the generally violent/antisocial perpetrators, were described as engaging in moderate to severe violence both within and outside of the family. Like dysphoric/borderline perpetrators, they are more likely to engage in psychological and sexual abuse in addition to physical abuse. They are also more likely to have an extensive criminal record and demonstrate psychopathology, including but not limited to psychopathy (Holtzworth-Munroe & Stuart, 1994). Perpetrators from the dysphoric/borderline and the generally violent/antisocial groups were identified as being violent more frequently and posing an ongoing risk for future assaults as compared to the relationship/family-only perpetrators (Stewart, 2001).

The 2000s

Holtzworth-Munroe and Stuart's (1994) typology gave rise to the idea that not all IPV perpetrators pose the same risk for recidivism, a finding that has been validated and replicated many times in the years following (Hamberger et al., 1996; Holtzworth-Munroe et al., 2000; Huss & Langhinrichsen-Rohling, 2006; Huss & Ralston, 2008; Lawson et al., 2003; Tweed & Dutton, 1998; Waltz et al., 2000). In the 2000s, various researchers found important differences in risk for recidivism and response to criminal justice intervention amongst the different types of IPV perpetrators. For example, Holtzworth-Munroe et al. (2003) found that 60% of the family-only men, 86% of dysphoric/borderline men, and 93% of generally violent men recidivated at least once during the three-year follow-up period. Huss and Ralston (2008) found that relationship/family-only IPV perpetrators are more likely to complete court-ordered treatment than the other two types of IPV perpetrators. They also found that compared to the other two types of IPV perpetrators, generally violent/antisocial perpetrators are most likely to recidivate, recidivate more frequently, and recidivate more quickly after completing a court-ordered intervention program. These findings provided additional support for the notion that a one-size-fits-all approach to policing IPV may not be the most efficient and appropriate use of criminal justice system resources, calling into question if interventions such as arrest or court-ordered treatments are equally effective for all subtypes of perpetrators (Stewart, 2001).

Although the debate about the value of pro-arrest and mandatory arrest policies from the 1990s continued into the 2000s, and important research was suggesting uniform approaches to policing IPV were misguided, these policies persisted in this decade. In

addition, new research was quickly demonstrating that the implementation of these policies was flawed. Studies of police response to IPV in jurisdictions with these types of policies demonstrated that policies alone do not ensure that all suspects are arrested (Belknap & Hartman, 2000; Buzawa, 1982; Ho, 2000). Belknap and Hartman (2000) conducted a study comparing different police jurisdictions to examine the effects of having pro-arrest policies. They found that even though pro-arrest policies were the only significant predictor of suspect arrest, there was still no arrest made in the majority of cases – even when a chargeable offense clearly occurred. Other research suggested that even if an IPV perpetrator was arrested, it was not guaranteed – or even likely – that they would be prosecuted (Sherman, 2000; Garner & Maxwell, 2009; Hartman & Belknap, 2003). For example, Garner and Maxwell (2009) examined the criminal justice system response and outcome in IPV cases and found that of all reported IPV cases, only about 5% resulted in prosecution and an even smaller proportion were ultimately convicted (Garner & Maxwell, 2009). Research also examined patterns in the execution of mandatory arrest policies, finding that police were most likely to arrest IPV suspects if the victim was a wealthy, White, older female (Avakame & Fyfe, 2001). Research also found that police under pro-arrest policies were least likely to arrest a suspect if they were White, if the assault took place in an urban centre, and if the victim had no visible injuries (Avakame & Fyfe, 2001; Leisenring, 2008).

Overall, there was no doubt that mandatory arrest policies were certainly influencing the rates of police intervention (e.g., in most jurisdictions the rates of arrests had increased by more than 70% since the implementation of the policies; Jaffe et al., 1993; Kendall-Tackett & Giacomoni, 2007; Nicholls et al., 2006); however, it was also

becoming increasingly salient that police officers were still acting as gatekeepers when deciding which IPV suspects would be arrested and charged (Cole & Gertz, 1998; Simpson et al., 2006). That is, police continued to make subjective judgments about which IPV victims and perpetrators would be the recipients of police intervention, regardless of policy contexts. This was argued to be a necessary process as the increased pressure on police to formally intervene for all IPV calls was not met with proportionate increases in resources (Storey et al., 2014). Police and other criminal justice resources (such as staffing, time, administrative support, availability of holding cells, availability of court dates, etc.) were lacking; therefore, resources needed to be directed to those IPV perpetrators with the highest risk for reoffending and highest intervention needs (Guo & Harstall, 2008; Roehl & Guertin, 2000; Trujillo & Ross, 2008; Viljoen et al., 2018).

Prior to the 2000s, police were largely engaging in unstructured risk assessment practices when responding to IPV, which involved making a subjective decision about an offender's dangerousness/risk for reoffending based on their expertise and experience (Guo & Harstall, 2008). The unstructured approach to risk assessment had been highly scrutinized for being inaccurate, unreliable, lacking transparency, and too easily influenced by personal biases (Helmus & Bourgon, 2011; Kropp, 2008; Nicholls et al., 2006). To help reduce subjectivity in the police decision-making process when responding to other types of crime (i.e., non-IPV crimes), evidence-based risk tools began to be incorporated in the risk assessment process (Bloom et al., 2005). The use of evidence-based risk tools offered a significant improvement over the unstructured approach in the accuracy of discriminating more dangerous offenders from those who

posed less risk for reoffending (Ægisdottir et al., 2006; Bloom et al., 2005; Litwack & Schlesinger, 1999; Northcott, 2012; Quinsey et al., 1998; Roehl & Guertin, 2000).

Risk assessment instruments have been used by Canadian police since the early 1990s for general and violent crime, but it was not until 2004 that the first evidence-based *IPV* risk assessment tool for police became available in Canada (despite *IPV*-specific risk tools being available to forensic experts and mental health professionals since the late 1990s; Bloom et al., 2005; Kropp & Hart, 2004). Having a tool designed specifically for police was an important distinction, as police typically do not have the training, resources, or information available to them to use the risk tools designed for mental health professionals; Hilton, Harris, & Rice, 2010). In 2008, Department of Justice officials conducted a study to identify which, if any, *IPV* risk tools were being used by criminal justice personnel across each province and territory of Canada (Millar, 2009). At that time, there were only four provinces/territories (i.e., Nova Scotia, New Brunswick, Ontario, and British Columbia) in which police officers were using at least one evidence-based *IPV* risk tool that was designed for police use. Police in two other provinces (i.e., Yukon and Alberta) were using at least one evidence-based *IPV* risk tool that was designed for use by mental health professionals rather than frontline officers. Police officers in another four provinces/territories (i.e., Newfoundland, Prince Edward Island, Québec, and Northwest Territories) were using risk tools or investigative checklists that were not evidence-based. In three provinces/territories (i.e., Manitoba, Saskatchewan, and Nunavut) no tools (evidence-based or otherwise) were being used to help police estimate risk for recidivism and guide their decision-making when responding to *IPV* (Millar, 2009). It is important to note that the use of evidence-based

risk tools varied tremendously by jurisdiction as no province/territory mandated the use of any given tool at the time of Millar's (2009) review.

Post-2010 (Current Context)

In the current Canadian context, police face a tension between fulfilling the government's responsibility to protect its citizens, sending a message that IPV will be taken seriously, and using discernment in managing police resources (Northcott, 2012; Viljoen et al., 2018). In addition, police are increasingly involved in risk management and violence prevention (Public Safety Canada, 2015). For instance, in 2019 the RCMP released a mandate stating that their role in responding to IPV is to enforce the law, engage in education and prevention measures, support the rehabilitation of offenders, and collaborate with other police agencies and social service providers to ensure victims are supported (RCMP, 2019).

The 2010s have seen an influx of research on the value of empirically-based IPV risk tools over and above subjective judgment, which has solidified its position as a key component of best practice in intelligence-led (i.e., evidence-based) policing in Canada (Canadian Association of Chiefs of Police [CACCP], 2016). However, the number of evidence-based IPV risk tools that are designed specifically for police officers is still limited. Additionally, provincial differences in which IPV risk instruments are selected for implementation can contribute to variation in how Canadian police officers respond to IPV calls (Department of Justice, 2015; Johnson & Dawson, 2011).

In 2013, the Department of Justice updated their study on which IPV risk tools were being used in Canadian provinces/territories (Millar et al., 2013). They partnered with the Public Health Agency of Canada's "Family Violence Initiative" department in

each province and territory and had them collect information about what risk assessment tools were being used by criminal justice personnel in their jurisdiction. They found that the number of provinces/territories in which police officers were using at least one evidence-based IPV risk tool specifically designed for police use had increased from four to six (i.e., with the addition of Alberta and Northwest Territories). Police in two provinces/territories continued to use evidenced-based IPV risk tools developed for mental health professionals, and one province/territory continued to use non-evidence-based tools/checklists. Nunavut and Saskatchewan were the only provinces/territories that reported having no tools available for police when responding to IPV (Millar et al., 2013). Although there is not a more recent update to this review of IPV risk tools used in Canada, it is certainly clear that evidence-based risk assessment tools are being increasingly endorsed by Canadian police agencies (Bourgon et al., 2018; Department of Justice, 2018b).

A qualitative study examining police officers' perceptions of the strengths and limitations of IPV risk assessment tools found that officers appreciate the consistency in data gathering, structure, and guidance for decision-making provided by structured risk assessment (Ballucci et al., 2017). However, some officers also report fearing that risk assessment tools could undermine police experience and discretion, that the process can be redundant with an experienced officer's routine investigation, and that it is a waste of time if the Crown or judge does not take the suspect's risk score into consideration. Another study by Campbell et al. (2018) surveyed 159 Canadian police officers from three different police agencies across New Brunswick about their experience with IPV risk tools. They found that 73.5% of officers find IPV risk tools at least somewhat

useful. Moreover, 67% of their respondents reported that they would use IPV risk tools if provided appropriate training (with younger age and increased desire for personal discretion in decision-making predicting officer's perception of lower usefulness of these tools; Campbell et al., 2018). These findings indicate that the majority of police officers in their sample were open-minded to using IPV risk assessment tools and could appreciate their value.

Although there is now plenty of literature and empirical research that emphasizes the need for tailored police responses to meet the unique needs of the victims and perpetrators of IPV, policies remain in effect in all Canadian jurisdictions to encourage arrest, charging, and prosecution in *all* cases of IPV (Burczycka, 2018; Department of Justice, 2015). In 2016, the Canadian Association of Chiefs of Police officially endorsed a National Framework for Collaborative Police Action to Intimate Partner Violence as an effort to inform and unify provincial policy, and thereby guide police practices across the country (Gill & Fitch, 2016). This framework outlines the expectations that police response to IPV should be supported by sound research, should target factors that create risk with effective case management strategies, and should prioritize harm reduction and IPV prevention (Gill & Fitch, 2016).

Despite increasing consistency across police services throughout Canada, there are still many variations in provincial and territorial criminal justice system policies that influence the roles, responsibilities, and discretion a police officer has when responding to IPV (Dawson & Hotton, 2014; Johnson & Dawson, 2011). The intervention and management of federal offenders is dealt with by Correctional Service Canada (CSC) at the federal level, but the majority of IPV perpetrators are dealt with at the provincial

level (Department of Justice, 2016). Provincial and territorial laws and policies are often developed in collaboration with grassroots and community-level organization, thereby reflecting the unique needs, resources, and culture of the different regions of Canada (Department of Justice, 2016). For instance, when police identify a high-risk IPV offender for whom there are not enough grounds for criminal charges in Manitoba, the victim involved qualifies for support through a provincial IPV victim support service. In Alberta, police risk assessment results are included in the offenders' court brief, and high-risk IPV offenders are referred to a specialized multi-disciplinary intervention team. In British Columbia, IPV cases identified by police as high-risk fall under a collaboration policy that requires timely sharing of risk assessment findings, court outcomes, breaches of conditions, and other relevant information about the offender and victim with all the key partners (e.g., police, Crown, correctional officers, victim services, and child protection workers; Department of Justice, 2016; McCuaig, 2012; Provincial Office of Domestic Violence, 2014). In New Brunswick, when a police officer issues an emergency order for an IPV perpetrator and deems the person to be at high-risk for continued IPV, a judge must review and confirm, deny, or vary the order within five days of submission to ensure timely processing of orders to protect victims of IPV (Legislative Assembly of New Brunswick, 2018).

Notable Oversights in the History of Policing of IPV

It is important to note that historically both public discourse and police services have overlooked and underserved certain groups, such as male victims of IPV (with female partners), IPV that occurs between same-gender couples, and IPV within Black, Indigenous, and People of Colour (BIPOC) communities. Some scholars have

speculated that the oversights regarding IPV victimization of cis/hetero-men and the LGBTQIA2S+ community was because the discourse was being driven by the feminist community, and these victims of IPV did not fit their conceptualization of IPV as a form of men's oppression of women (McLaughlin & Rozee, 2001; Ristock, 2001, 2003; Rollè et al., 2018). The underrepresentation of male victims (with female partners) also may have been influenced by the culturally-sanctioned masculine ideals in Canada, creating a hesitancy in men to seek help (from police, other professionals, or within their personal networks) when experiencing IPV, which continues to be a barrier currently (Ansara & Hindin, 2010; Burczycka, 2016; Cotter, 2018; Cotter & Savage, 2019; Lysova et al. 2020; Migliaccio, 2001, 2002). Likewise, historically there has been (and continues to be) a silence surrounding violence in the LGBTQIA2S+ community (Rollè et al., 2018).

Although the literature on police response to IPV calls involving complainants who are male (with female partners) or in same-sex intimate partner relationships is scarce, there are some studies that have examined this issue. A commonality between both groups is that there is often a hesitancy to involve police out of a fear that it will not help and that it may make their situation worse. For instance, several studies have demonstrated that people from the LGBTQIA2S+ community historically have reported experiencing additional victimization, homophobia, or disbelief from police officers after reporting IPV (Barnes, 1998; Burke et al., 2002; Pattavina et al., 2007; Guadalupe-Diaz & Yglesias, 2013). Likewise, for male IPV victims (with female partners), there is often also a fear of being confused as the perpetrator when police arrive on the scene (Roebuck et al., 2020). One Canadian study found that 64% of male victims of IPV who called the police were treated as the perpetrator (Dutton, 2012), and another study found

that male victims of IPV were four times less likely to have a restraining order issued when they called the police for IPV (Lysova et al., 2019). Current definitions of IPV used by Canadian police have been expanded to include female-to-male and same-sex IPV, but recent research largely suggests that heteronormative understandings of IPV continue to influence police responses to IPV (Guadalupe-Diaz, 2016; Rollè et al., 2018; Whitehead et al., 2021).

The policing of IPV within BIPOC communities is another subject that has been too often overlooked in related academic discourse, and one that ought to be situated in the historical context of racialized policing in Canada. A rich scholarship on the history of racialized policing exists from BIPOC Canadian authors such as Robyn Maynard, Elizabeth Comack, Barrington Walker, Peggy Bristow, Dorothy Williams, Tiffany King, and many others. Although it is beyond the scope of the current dissertation to do a thorough review of this literature, works by these authors are recommended to contextualize the storied and often tense relationships between police officers and the visible minority suspects, witnesses, and complainants they encounter. This history is rooted in both colonization and slavery, which normalized the surveillance of Black and Indigenous people through tracking down runaway slaves, restraining Indigenous adults to small reserves of land, and confining Indigenous children within residential schools (Comack, 2012; Maynard, 2017; Owusu-Bempah, 2017). The heightened surveillance of BIPOC individuals, and the over-policing of the neighbourhoods and communities where they reside, continues today and contributes to BIPOC individuals experiencing disproportionate rates of arrest, charges, prosecution, conviction, and incarceration (Comack, 2012; Maynard, 2017; Kubik et al., 2009). For instance, Black and Indigenous

Canadians are approximately three and five times more likely (respectively) to be incarcerated for the same crimes as White Canadians (Sapers, 2015; Wortley & Owusu-Bempah, 2009; Zinger, 2021).

BIPOC individuals are also over-represented as the victims of general violence and IPV, especially Indigenous women (Brownridge, 2008; CACP, 2016; Daoud et al., 2012; Daoud et al., 2013; Department of Justice, 2017b). For instance, a study from nearly three decades ago, in which 621 Indigenous people were interviewed (i.e., just two years before the last residential school in Canada closed; Government of Northwest Territories et al., 2013), La Prairie (1995) found that as much as 70% of male and 75% of female Indigenous Canadians reported experiencing family violence in their childhood. More recently, national self-report and official record data from 2015 indicated that Indigenous women experience higher rates of IPV, more severe injuries, and are six times more likely to be killed by their intimate partner compared to non-Indigenous women (Department of Justice, 2017b). Likewise, national self-report data from 2018 found that 61% of Indigenous women compared to 44% of non-Indigenous, non-visible minority women reported experiencing some form of IPV in their lifetime (Cotter, 2021). These higher rates of violence experienced within BIPOC communities further represents the legacy of colonialism, slavery, racial discrimination, assimilation, and exclusion (Varcoe et al., 2017; CACP, 2016; Rotenberg, 2019).

Given the longstanding tension between police services and the BIPOC community, it is perhaps no surprise that research has shown that visible-minority individuals are less likely than Caucasian individuals to report most crimes to police (Frank et al. 2005; Gabbidon & Higgins, 2009; Guadalupe-Diaz, 2016; Lai & Zhao,

2010). On the other hand, when it comes to IPV-related crimes specifically, previous research spanning several decades has demonstrated that BIPOC women are *more* likely than their White counterparts to call the police as their primary form of help-seeking (Guadalupe-Diaz, 2016; Kaukinen, 2004; Krishnan et al., 2001; McFarlane et al., 1997; Pearlman et al., 2003; West et al., 1998). This phenomenon is not well understood, but some have speculated that it may be due to being desensitized to police presence in their neighbourhoods and having no or limited access to other resources in the face of IPV (Guadalupe-Diaz, 2016). Despite calling the police for help, BIPOC victims of IPV report much stronger negative perceptions of the police and fear of the police than White IPV victims, as well as being significantly *less* likely to rate police as being good at enforcing the law, increasing their safety, prompt in their response to the call, approachable, and fair (Barret et al., 2019; Guadalupe-Diaz, 2016; Maynard, 2017; Ibrahim; 2020). The fact that the Canadian government has been critiqued for failing to adequately address the startlingly high rates of missing and murdered Indigenous women in Canada illustrates that police intervention has not always led to increased victim safety for BIPOC Canadians (Amnesty International, 2014; Barrett & Peirone, 2019). Overall, these findings offer a reminder that victims' experiences of IPV and police intervention throughout history need to be considered through an intersectional lens.

Summary of Historical Context

The role of police officers in responding to IPV has changed greatly over the last 50+ years in Canada. While police in the 1960s and 1970s were discouraged to make arrests or seek charges for IPV, the 1980s saw an increased scrutiny of police response to IPV fueled by the feminist movement. Unfortunately, this movement largely

overlooked and underserved those IPV victims who were male (with female partners), in the LGBTQIA2S+ community, and BIPOC individuals. Mandatory arrest policies in the 1990s resulted in a substantial increase in police intervention in IPV situations, but it was clear that police were still using discretion and subjective judgments of risk to decide when and for whom to intervene. Theory, feminist discourse, and research in the 2000s led to an increased understanding that it is inappropriate and ineffective for the police to have a uniform response to every IPV situation. In 2004, the first police-IPV risk assessment instrument was released to aid police officers in making judgements of risk for IPV recidivism, and since the 2010s risk tools have become a common part of policing IPV. Current definitions of IPV used by police have been expanded to include female-to-male and same-sex IPV, but there is a continued need to deconstruct the heteronormative biases around IPV to improve police response for these groups. In the current Canadian context, the role of the police is influenced by the province/territory and jurisdiction in which they operate, but Canadian police are expected to be actively involved in risk management and the prevention of IPV (Gill & Fitch, 2016). Evidence-based risk assessment is currently understood as a key part of risk management and IPV prevention, the conceptualization and application of which has evolved substantially over time.

The Evolution of General Risk Assessment

First-Generation Risk Assessment

Long before IPV-specific risk tools became available, the process of estimating the likelihood that an offender will recidivate (i.e., risk appraisal or assessment) in order to appropriately manage offenders and direct resources has played an important role in

the broader criminal justice system. In the Canadian context, however, risk assessment was originally used in psychiatric hospitals (Bloom et al., 2005). The first formal consideration of “dangerousness” in Canada was under the direction of Dr. Kenneth Gray at the Toronto Psychiatric Hospital in the 1940s (Bloom et al., 2005). Gray used his experience working with difficult and dangerous patients to publish advice for hospital staff about characteristics of aggressive and violent patients, and to offer sensible and practical suggestions about the safe management of dangerous patients (Fidler & Gray, 1947). This approach to risk assessment, which involved subjective judgment based on training and experience, is referred to as “unstructured professional judgment.” The strength of this unstructured approach to risk assessment was that it was client-centered and flexible enough to allow for context-specific tailoring of risk management strategies. It was predominant for much of the twentieth century and became known as the first-generation of risk assessment (Bonta & Andrews, 2007). However, the general field of risk assessment has progressed a great deal, both empirically and theoretically, since this time when unstructured professional judgment was considered best practice.

The accuracy of unstructured professional judgment came into question in the 1970s. For example, after a US Supreme Court Ruling deemed that mentally ill offenders were being detained for inappropriately long sentences, hundreds of mentally ill offenders who had previously been considered “high-risk” and dangerous were released from maximum security institutions (Steadman & Coccozza, 1974). Researchers followed these patients for four years following their release, finding that only about 20% of them were reconvicted (Steadman & Coccozza, 1974). Just a few years later, a

similar study followed hundreds of patients who were also deemed high-risk offenders through risk assessments of unstructured professional judgment. After a three-year follow-up, it was found that only 14% had violently recidivated (Thornberry & Jacoby, 1979). In a well-known review of the literature on risk prediction, Monahan (1981) concluded that mental health professionals were only able to predict violent recidivism in less than one-third of cases when using this informal risk assessment approach. There was a growing recognition that the unstructured professional judgment approach to risk assessment lacked predictive accuracy and had poor interrater reliability due to its subjective nature.

Second-Generation Risk Assessment

In response to the limitations of first-generation risk assessment, the second-generation of risk assessment emerged with standardized, evidence-based risk tools. These tools are actuarial in nature, with quantitative risk estimates based on empirically-identified risk factors that are often statistically weighted (Bonta & Andrews, 2007). These risk factors are typically static/historical in nature, such as having a history of criminal behaviour and history of violation of court orders. Item scores are summed and produce a “risk score,” with higher scores representing a higher risk that an offender will commit a specified type of crime, within a specified time period (Kropp & Hart, 2004). These risk tools offer improved accuracy in risk estimation and almost always outperform unstructured clinical judgment (Ægisdottir et al, 2006; Grove & Meehl, 1996). As a result, many correctional jurisdictions began adopting the use of risk assessment tools in the 1980s (Bonta & Andrews, 2007).

A well-known example of a second-generation, actuarial risk tool is called the Violence Risk Appraisal Guide (VRAG; Harris et al., 1993). The VRAG was developed after psychologist Vernon Quinsey was appointed in 1975 to create a research department at Oak Ridge, a maximum-security psychiatric hospital for offenders with severe mental illness in Ontario, Canada (Bloom et al., 2005). Like other research of the 1970s, Quinsey found that clinicians' estimates of risk were no more valid than non-professionals, and clinicians from different disciplines tended not to agree with each other about relevant risk factors and risk management (Quinsey & Ambtman, 1979). In response, Quinsey put together a team of professionals in Ontario who joined forces to review the empirical literature on violent recidivism and to statistically identify risk factors for future violent behaviour (Hilton, Harris, & Rice, 2010; Northcott, 2012). The goal of their research was to develop a checklist of offender characteristics that could be used to predict violent behaviour in forensic populations (Hilton, Harris, & Rice, 2010). From this effort came the VRAG (Harris et al., 1993), which predicted violent recidivism with a large effect size and has garnered a vast amount of research support over time (Campbell, French, & Gendreau, 2009; Hanson & Morton-Bourgon, 2007; Quinsey et al., 2006).

Second-generation risk instruments like the VRAG are still commonly used in Canada, but they too have been critiqued. One of the major limitations of second-generation actuarial risk assessments is the fact that they are atheoretical (i.e., the risk items were strictly chosen due to statistical significance, not theoretical significance or rehabilitative value; Bonta & Andrews, 2007). The result is that actuarial risk tools are predominantly comprised of static (i.e., historical) risk items. Static risk items are not

able to account for changes in risk level over time, which limits their utility for risk management and risk reduction (Hoge & Andrews, 1996). Critics of the second-generation risk tools suggest that risk assessment with the sole goal of risk prediction is incomplete, stating that the true goal of risk assessment should be to inform risk management and reduction.

Third-Generation Risk Assessment

The RNR model of offender rehabilitation became one of the most influential theories on third-generation risk assessment, with an explicit emphasis on the assessment of both overall risk level and criminogenic needs to inform appropriate intervention targets and intensity (Andrews & Bonta, 2006; Bonta & Andrews, 2017). The inclusion of dynamic risk items became a hallmark of the third-generation of risk assessment (Holtzworth-Munroe & Stuart, 1994; Kropp & Hart, 2004). Dynamic risk factors reflect internal states or temporary circumstances that change over time, such as employment and financial problems, relationship problems, and problematic substance use (Northcott, 2012). Given that dynamic risk factors can change with time or treatment, the advantage of third-generation risk tools is that they are sensitive to fluctuations in risk level (Andrews & Bonta, 2003; Heilbrun, 1997). Most of these third-generation tools are still actuarial in nature, with a final risk score that categorizes the offender as high-, moderate-, or low-risk for recidivism. However, it is worth noting that an alternative type of risk tool emerged around this time in response to the critique that highly mechanical, actuarial tools lack the flexibility to consider the unique context of individual cases (Bonta, 2002; Borum, 1996). This alternative approach was called “structured professional judgment (SPJ).” The SPJ approach to risk assessment is still

based on empirically and theoretically relevant risk factors, but a risk score is not the final determinant of risk level. Instead, assessors consider all the risk factors and make a decision about risk level based on this analysis, whilst considering the individuals involved and the broader context of the case (Kropp & Hart, 2004).

An example of a third-generation risk tool is the Historical, Clinical, and Risk Management Scheme-20 (HCR-20; Webster et al., 1997). It was developed by researchers and mental health professionals in British Columbia in response to the 1992 federal law Bill C-30, which forbids the automatic and indeterminate detainment of those who were found not criminally responsible on account of mental disorder (Bloom et al., 2005; *Criminal Code*, 1985, c C-46, s 16(1)). This new law specified that only those who posed a “significant threat” to public safety could be detained, making risk assessment all the more a crucial part of the criminal justice system. Many forensic and mental health clinicians in British Columbia collaborated and consulted to identify empirically-based risk factors that were also theoretically relevant and had rehabilitative significance (Bloom et al., 2005). The eventual outcome was the HCR-20, which was designed to be sensitive to change in risk over time (Belfrage & Strand, 2001; Brunero & Lamont, 2009; Dolan et al., 2002). The HCR-20, which has a great deal of research support, includes ten historical/static risk items and ten dynamic risk items (made up of five “criminogenic needs” items and five “risk management” items; Webster et al., 1997). Its latest version has expanded SPJ decisions to not only address whether the person will reoffend violently, but also the imminence of that risk and the likelihood of harm to the victim should it occur.

Fourth-Generation Risk Assessment

The fourth and most recent generation of risk instruments build on previous generations' tools by interweaving the risk assessment process with intervention, as a means of identifying treatment targets, general treatment considerations, areas of success, and changes in risk over time (Bonta & Andrews, 2007). This approach to risk assessment recognizes that not only does risk change, but treatment needs change as well (Andrews et al., 2006). Therefore, fourth-generation risk instruments are administered frequently during the course of offender case management and treatment so positive changes can be identified and reinforced, and areas requiring intervention modification can be identified quickly in order to maximize potential risk reduction (Andrews & Bonta, 2003; Campbell, French, & Gendreau, 2009). An example of a fourth-generation risk assessment is the Level of Service/Case Management Inventory (LS/CMI; Andrews et al., 2004). The LS/CMI was designed to assist criminal justice professionals with decisions about the management, supervision, and case planning for adult offenders. Key areas measured include criminal history, education, employment, family/marital issues, leisure and recreation, companions, alcohol and drug problems, attitudes, criminogenic needs, responsivity, and case management factors (Andrews et al., 2004). A meta-analysis on the predictive validity of the LS/CMI indicates that its total score and criminogenic needs domain (i.e., dynamic risk) have moderate predictivity of future general and violent recidivism (Olver, Stockdale, & Wormith, 2014).

A meta-analysis by Campbell, French, and Gendreau (2009) examined the predictive validity of risk tools from various generations for violent recidivism. Their meta-analysis included 88 studies and over 70 different risk measures in total. Their

results indicated that third-generation risk tools demonstrated a mean effect size ($Z^+ = .25$) higher than that of second-generation tools ($Z^+ = .22$) for predicting violence, except in institutional settings (where second-generation tools are somewhat better at predicting violent behaviour). Overall, these authors concluded that both second- and third-generation risk instruments are able to predict violence recidivism with at least moderate success. Similarly, Gendreau et al. (1996) conducted a meta-analysis examining over 1000 correlations between many predictor variables and violent recidivism. They found that dynamic risk factors/criminogenic needs had significantly higher mean r values (.15) than static factors (.12). As the risk assessment process has evolved over the years, it is evident that the addition of dynamic risk factors in the third- and fourth-generation tools has been an important development in the field theoretically and practically.

Dynamic Risk Factors

Dynamic risk factors represent the distinction between “risk status” and “risk state.” Risk status is based on static risk factors that are immutable over time, which is used to compare level of risk from one person to the next (i.e., inter-individual risk; Klepfisz et al., 2016). Risk state, on the other hand, is based on dynamic risk factors and considers psychological, situational, social, and biological factors of a particular person, at a particular time (i.e., intra-individual risk; Douglas & Skeem, 2005; Klepfisz et al., 2016). To further conceptualize dynamic risk factors, Douglas and Skeem (2005) argued that there are three important components to dynamic risk factors:

1. the factor must precede, and increase the propensity for, violence (i.e., a risk factor);

2. the factor must be changeable, either spontaneously or due to intervention effects (i.e., a dynamic factor); and
3. the likelihood of violence must change when the factor changes (i.e., a causal factor).

A limitation in the current literature is that these three criteria for dynamic risk factors are not given enough attention (Cording et al., 2016; Klepfisz et al., 2016). For example, Klepfisz et al. (2016) suggest it is difficult to establish whether a factor truly precedes and increases violence without expensive, longitudinal studies. As a result, there is not a robust literature to support identified dynamic risk factors as truly causal factors. Cording et al. (2016) stress that it is not enough for research to test the hypothesis that dynamic risk factors are *actually* changeable (i.e., criterion two), but it must also examine the impact of fluctuations in dynamic risk factors on recidivism outcomes (i.e., criterion three; Cording et al., 2016). Research has been slowly accumulating to address Douglas and Skeem's (2005) three criteria, as reviewed below.

Dynamic Risk Factors for Offending (Criterion One)

The developers of the RNR model have identified seven areas of dynamic risk that have been well established as predictors of general offending by many researchers (Andrews & Bonta, 2003). They include antisocial personality pattern (i.e., impulsive, aggressive, manipulative, risk-taking, exploitative), antisocial cognition (i.e., values, beliefs, and ways of thinking that favour and reinforce criminal behaviour), antisocial peers (i.e., preferring to associate with peers to reinforce criminal behaviour), family/relationship problems (i.e., chaotic and poor quality relationships), school/work problems (i.e., poor performance and/or limited engagement), leisure and recreation

problems (i.e., lack of involvement in prosocial activities), and substance use problems (i.e., problematic use of alcohol and/or drugs; Andrews & Bonta, 2003). The RNR model does not preclude the examination of other risk factors, dynamic or otherwise, but rather suggests that these seven factors (plus history of antisocial behaviour, a static risk factor) are most robustly related to general reoffending and, therefore, should be considered in the assessment of risk (Andrews et al., 2011).

Although the literature on dynamic risk factors for general offending is vast, the research on dynamic risk factors for IPV recidivism specifically is still emerging. However, there is sufficient research to draw from to identify several promising dynamic risk factors for IPV. To do this, an electronic literature search was conducted using the University of New Brunswick's electronic library database, called "UNB WorldCat." Searches through UNB WorldCat included sub-searches through over thirty journal databases (e.g., PsycArticles, SAGE Journals, SpringerLink). Key search terms included (a) assessment-related terms (e.g., actuarial, clinical, prediction, ODARA, SARA, B-SAFER), (b) risk factor terminology (e.g., risk, likelihood, probability, dynamic, static, historical, acute, stable), (c) terms related to IPV (e.g., intimate partner violence, domestic violence, spousal assault, intimate partner abuse, partner battering), and (d) terms related to outcomes (e.g., recidivism, reoffending, reoccurrence, IPV cycle of violence, desistance). Additional studies were identified by reviewing reference sections of relevant articles and the manuals of existing IPV risk tools. Variables that are conceivably changeable/dynamic and that the literature consistently identified as a significant predictor or correlate of IPV were considered potential dynamic risk factors. The identified risk factors include perpetrator state factors (i.e., jealousy/control,

antisocial attitudes, substance use, and self-harm/suicidal ideation) and contextual factors (i.e., relationship problems, escalation in IPV frequency or intensity, employment or financial problems, and victim fear of future IPV).

Perpetrator State Factors. In terms of perpetrator state factors, the IPV literature has reliably indicated that excessive jealousy/control, antisocial attitudes, substance use, and self-harm/suicidal ideation are important dynamic factors that contribute to the prediction of future IPV behaviour. In terms of jealousy/control, research has examined this concept from the perspective of controlling behaviours, possessiveness, unprovoked accusations of infidelity, isolation from friends and family, obsessively checking in, and other forms of psychological abuse that serve to humiliate, harm, exploit, intimidate, dominate, and isolate their partner (Hilton et al., 2001; Stark, 2009). Cattaneo and Goodman (2003) conducted a prospective study with 169 female victims of IPV, finding that perpetrator efforts to control all or most of their partner's daily activities and associations is a particularly strong predictor of future physical assault. Others have found that a sense of ownership or possessiveness (e.g., "if I can't have her, no one will") is significantly correlated with severe physical IPV (Davies et al., 1998; Hart, 1990). In addition, research on intimate partner homicide has reliably found that excessive jealousy, perceived infidelity, and control of an intimate partner are amongst the most common motivators of intimate partner homicide (Elisha et al., 2010; Hannah et al., 1998; Hanzlick & Koponen, 1994).

Research on antisocial attitudes involves examining the impact of political, religious, cultural, and personal attitudes that condone or support the use of violence or other criminal behaviour. People who hold such attitudes also tend to deny the authority

of the law and police officers who enforce it², which is associated with behaviours such as violations of court or police orders (Kropp & Hart, 2004; Moser, 2012). Several studies have found that attitudes that condone the use of violence are significantly correlated with IPV (Dutton & Kropp, 2000; Eriksson & Mazerolle, 2015; Robertson & Murachver, 2009). For example, O’Keefe (1998) found that the risk of IPV in dating relationships increases when an individual believes that the use of violence is an appropriate and justifiable way to resolve conflict. Others have found that minimization or denial of responsibility in IPV is significantly correlated with future IPV behaviour (Grann & Wedin, 2002). In a study by Petersson et al. (2016) comparing types of IPV perpetrators, the presence of attitudes condoning violence was found to be one of the most informative acute risk factors for IPV recidivism among offenders whose violence was limited to the family only. Another study that used hypothetical vignettes about marital problems demonstrated that spouses who had committed IPV were more likely than non-violent spouses to conclude that their partner had acted with malicious intent in the vignette (i.e., that they were being deliberately insulting or hurtful; Holtzworth-

² At the time of proposing this dissertation and collecting, analyzing, and interpreting the data, it was commonly accepted to include “anti-authority behaviours or beliefs” (e.g., opposed to police presence or involvement, aggressive towards law enforcement, etc.) as a part of the operationalization of “antisocial attitudes” (Andrews, 2015; Andrews & Bonta, 2006; Bonta & Andrews, 2007; McDougall, 2014; Moser, 2012). Such an operationalization is even included in various measures of risk for criminal recidivism (Cooke, 1996; Mills et al., 2004). Given this common approach, some of these behaviours were included in the operationalization of the ‘antisocial attitudes’ variable considered in this study (described in the “Methods” section). In doing a final read-through of my dissertation prior to submitting to my defence committee, and in the midst of a changing cultural conversation about racialized policing, I came to see this variable in a different way. What I previously thought of as an intuitive indicator of antisocial attitudes I now regard with caution and apprehension as these behaviours could potentially stem from a suspect’s experience with racism. That is, these anti-authority behaviours and beliefs *could* represent true ‘antisocial attitudes’ held by a suspect. However, for individuals who are Black, Indigenous, and People of Colour (BIPOC), these indicators also could represent a reaction of distrust towards police due to historical and current issues of racial profiling and mistreatment of BIPOC individuals by the police (Ibrahim; 2020; Statistics Canada, 2020b). In recognizing this as a possible source of bias in my dissertation, I have used various footnotes throughout to provide alternative analyses and interpretations when possible. This issue is also further discussed in an addendum after the “Implications and Conclusions” section.

Munroe & Hutchinson, 1993). They tended to at least partially blame the IPV victim for their experience of abuse, thereby justifying the behaviour. Research also has found that high levels of minimization, denial, victim blaming, and justification are associated with decreased IPV perpetrator treatment compliance in addition to increased risk for IPV recidivism (Daly & Pelowski, 2000; Henning et al., 2005; Kropp et al., 1995).

The research on substance use and IPV is extensive, demonstrating a strong link between alcohol and illicit drug use and IPV in a variety of samples (Campbell & Glass, 2009; Gondolf, 2002; Moore et al., 2011; Saunders, 1995; Schumacher et al., 2001). This relationship is maintained even after controlling for other known risk factors, such as relationship problems and antisocial attitudes (Leonard & Senchak, 1996). Gondolf (2002) conducted a large-scale, four-year study to identify risk markers for IPV reoffending. He found the most influential risk marker was excessive alcohol use, which was associated with a 3.5 times increase in IPV recidivism compared to those who did not become intoxicated. Those perpetrators who used excessive alcohol nearly every day were 16 times more likely to reoffend compared to those who seldomly or never drank alcohol. One prospective daily record study found that IPV was about 10 times more likely on drinking days, and increased alcohol consumption was associated with increased severity of IPV (Fals-Stewart, 2003). They also found that 60% of IPV reoffending in their sample occurred within two hours of drinking, and 80% occurred within four hours (Fals-Stewart, 2003). Several prospective studies have replicated these findings with other samples (Murphy et al., 2005; Parks et al., 2008; Moore et al., 2011).

Various other studies have examined the relationship between IPV and illicit drug use in particular (Campbell et al., 2003; Fals-Stewart et al., 2002; Herbert et al.,

2021; Klostermann, 2006; Yu et al., 2019). For instance, Campbell et al. (2003) interviewed 563 IPV victims or their proxies to explore risk factors for intimate partner homicide. They found that illicit drug use was as relevant in predicting intimate partner homicide as access to a weapon, and it was a better predictor than alcohol use (Campbell et al., 2003). A population-based, longitudinal study from Sweden examined the association between nine diagnostic groups of mental disorders and IPV, finding that individuals with a drug use disorder had the highest rates of IPV against their partners (Yu et al., 2019). In addition, Fals-Stewart et al. (2002) found that of those IPV perpetrators with a history of illicit drug use, those who received targeted treatment were less likely to commit a new IPV act in the following year.

Another consistent finding in the literature on dynamic risk for IPV is a strong relationship between suicidal thoughts and IPV behaviour. One self-report study, using a sample of college students, found that IPV perpetration was significantly correlated with higher levels of suicidal ideation (Lamis et al., 2013). Other studies have consistently found a moderately strong relationship between threats, thoughts, and attempts of suicide and IPV perpetration (Blosnich & Bossarte, 2012; Coker, Smith, et al., 2002; Flynn et al., 2016; Hilton et al., 2001). In addition, threats of self-harm have been found to be a significant predictor of future IPV (Hart, 1990).

Contextual Factors. Four key contextual factors have been identified as dynamic risk factors for IPV: relationship problems, escalation of abuse, employment/financial problems, and victim fear. In terms of relationship problems, research has reliably found that physical separation, anticipated separation, or starting the legal separation process increases risk for intimate partner homicide (Belfrage &

Rying, 2004; Campbell, 2012; Campbell & Glass, 2009; Campbell et al., 2003; Capaldi et al., 2012; Dobash et al., 2004). Likewise, one study found that a woman's expressed desire to leave her romantic partner was amongst the leading motivators of intimate partner homicide (Elisha et al., 2010). Another study has found that frequent conflict and on/off relationship cycles is significantly correlated with severe IPV (Kropp, 2008). Various other studies have found a strong correlation between increased severity of marital discord/relationship distress with increased severity of IPV (Hanson et al., 1997; Harris et al., 1993; Pan et al., 1994; Rosenbaum & O'Leary, 1981). Based on the analysis of data from many countries and samples, it has been theorized that IPV in the face of relationship problems and separation is the result of the threat of loss of control over the relationship (i.e., IPV as an effort to regain control of the relationship; Belfrage & Rying, 2004; Elisha et al., 2010; Sheehan et al., 2015; Wilson et al., 1995).

Various studies have found that an escalation in violent threats, verbal abuse, and other abusive behaviours is a robust predictor of imminent IPV recidivism and IPV lethality (Garcia et al., 2007; Petersson et al., 2016; Sheehan et al., 2015). This relationship has been reliably demonstrated whether escalation is defined in terms of increase in severity (i.e., relative seriousness of impact or potential impact of IPV), increase in frequency of IPV, or both (Campbell, 1986; Davies et al., 1998; Hilton et al., 2001). Previous research has also consistently found a relationship between employment/financial problems and risk for IPV, such as unstable employment, unemployment, and low income (Dutton & Kropp, 2000; Moffitt et al., 1997; Saunders, 1995). For example, Hanson et al. (1997) found that the most severe IPV was predicted by the lowest income level, the highest relationship problems, and the most severe

depression symptoms. A comprehensive study by Campbell et al. (2003) examining predictors of intimate partner homicide found that the perpetrator's unemployment was the strongest sociodemographic risk factor for intimate partner homicide.

Finally, previous research has consistently identified one victim factor that is a robust predictor of IPV recidivism: victim fear (Cattaneo & Goodman, 2003; Goodman et al., 2000; Heckert & Gondolf, 2004). Victimized partners can be highly accurate at predicting the likelihood of future violence because they know their partners better than anyone else and are especially attuned to their cycles of violence (Hart, 1994). Weisz et al. (2000) found that female victims' predictions of future IPV provided incremental predictive validity over and above a variety of other risk factors, such as previous choking, rape, threats to kill, and sexual jealousy. In a four-year, multisite study, Gondolf (2002) found that victims' perceptions of the likelihood of re-assault emerged as the strongest and most consistent risk factor for recidivism, emerging as a better predictor than all of the perpetrator characteristics combined.

Changeability of Risk Factors (Criterion Two)

There do not appear to be any studies using samples of IPV perpetrators that evaluate Douglas and Skeem's (2005) second criterion for dynamic risk factors by examining whether or not they fluctuate over time. There is, however, research using samples of sexual offenders and general adolescent offenders that examine this second criterion. For example, Hudson et al. (2002) found significant changes in pre-post treatment dynamic risk scores (e.g., trait anger, suppressed anger, perspective-taking ability) among a sample of 219 adult male sex offenders. Similarly, Brooks-Holliday et al. (2012) examined a variety of risk factors in a sample of 88 male offenders from a

forensic treatment facility. They measured participant risk factors within one week of admittance and within one week of release. They found significant improvements between pre- and post-treatment dynamic risk scores (e.g., pro-criminal attitudes, marital relationships, education, and employment). Another example includes a study done by Viljoen et al. (2012), who prospectively followed 90 adolescent offenders (62 male, 28 female) for a three-month period. They found that 92% of their sample showed either increased or decreased dynamic risk. All these studies contribute to the literature supporting the notion that these dynamic risk factors are, in fact, changeable and thereby meet Douglas and Skeem's (2005) second criterion. However, there is a need for research to examine the changeability of IPV-specific dynamic risk factors using a sample of IPV perpetrators.

Impact on Recidivism (Criterion Three)

Like the previous criterion, there does not appear to be any IPV research evaluating the effect of change in dynamic risk on recidivism outcomes, Douglas and Skeem's (2005) third criterion. Only a limited number of studies have examined criterion three using sexual and general offender samples, with mixed findings. For example, Beggs and Grace (2011) examined the impact of dynamic risk changes on recidivism with a sample of 218 adult males who sexually offended against children. Participants completed a prison-based cognitive-behavioral treatment program that targeted dynamic risk factors. Their results showed that changes in dynamic risk factors in the desired direction were associated with reduced sexual recidivism, with an average follow-up period of 12.24 years. Likewise, Cohen et al. (2016) followed more than 64,000 offenders on federal supervision in the United States. They tracked recidivism,

static risk, and dynamic risk. These risk scores were calculated at the time of the index offence and 12 months after first re-offence. Cohen et al. (2016) found that high-, moderate-, and low-risk offenders who had decreases in their dynamic risk score were less likely to reoffend compared to counterparts whose dynamic risk levels were unchanged or increased. They also found that increases in dynamic risk classification and/or score was associated with increased rates of re-arrest. Not surprisingly, their results demonstrated that low-risk offenders whose scores decreased saw no change in recidivism (i.e., a floor effect was taking place). Several other studies have found similar results, with changes in dynamic risk scores reflecting parallel changes in recidivism (de Vries Robbé et al., 2015; Labrecque et al., 2014; Lewis et al., 2013; Olver, Nicholaichuk, et al., 2014; Vose et al., 2013; Wilson et al., 2013).

Despite some support for a link between changes in dynamic risk and recidivism outcomes, several studies have found no significant association between changes in dynamic risk scores and recidivism (Barnett et al., 2013; Goodman-Delahunty & O'Brien, 2014; Hanson, Harris, et al., 2007; Viljoen et al., 2016; Viljoen et al., 2017). For example, Barnett et al. (2013) employed a sample of 3402 convicted adult sexual offenders attending probation treatment. The average treatment length was 14 months, and the average follow-up period was three years. They found that the degree of change in two of the three dynamic risk categories examined (i.e., pro-criminal attitudes and self-management skills) was not associated with reductions in sexual and/or violent recidivism. Only change in a third dynamic risk category, socio-affective functioning, was associated with change in recidivism in their sample (Barnett et al., 2013). Viljoen et al. (2017) found comparable results with a sample of 163 adolescent sexual offenders

who attended a residential, cognitive-behavioural treatment. The average time in treatment was 1.13 years. Despite the fact that about half of the sample demonstrated reliable decreases in dynamic risk, reductions in dynamic risk were not associated with reductions in recidivism (Viljoen et al., 2017). In addition, post-treatment risk ratings were not any more accurate at predicting recidivism than the ratings conducted pre-treatment. Viljoen et al. conducted a similar study in 2016 with a sample of 156 youth on probation for general and violent crimes (i.e., non-sexual crimes). These youth were prospectively followed for two years, with reassessment of risk every three months for the first year. Their results also demonstrated a lack of association between change in dynamic risk and changes in recidivism (Viljoen et al., 2016).

Viljoen et al. (2016, 2017) have speculated about why some research has failed to find an association between changes in dynamic risk and recidivism, such as methodological differences and measurement error. They also suggested that changes in recidivism may not be observed in research with longer follow-up periods because dynamic risk levels may have continued to fluctuate over time (i.e., treatment effects may not have been sustained over the long-term). Viljoen et al. (2016) also discussed the possible effect of type of dynamic risk factor being measured. That is, some dynamic risk factors may change more gradually over time, and some may change rapidly in response to immediate circumstances. In fact, dynamic risk factors can be classified into two subtypes: stable and acute dynamic risk factors. Stable dynamic risk factors for violence could include skills, deficits, behavioural patterns, personality, self-management abilities (i.e., impulse control), positive beliefs about violence, and poor attachment. Stable dynamic risk factors can change slowly over months or years, and

change is facilitated by intentional, effortful intervention (Hanson & Harris, 2000, 2001; Hanson, Harris, et al., 2007). Acute dynamic risk factors, on the other hand, can change spontaneously or with intervention/effort in a matter of weeks, days, or even hours (e.g., intoxication, active psychosis, and angry mood; Hanson & Harris, 2000, 2001; Hanson, Harris, et al., 2007). Some dynamic risk factors can be considered both acute and stable. For example, if a person experienced problematic substance use for many months or years, it would be considered a stable dynamic risk factor for future violence. If they were also intoxicated at the time that they committed the offence, substance use would also be considered an acute risk factor.

Dynamic Risk Prediction

Overall, there are several mixed findings and gaps in the research that seeks to identify true dynamic risk factors as per Douglas and Skeem's (2005) criteria, and most of this research has been conducted with sexual offenders (Yesberg & Polaschek, 2015). However, research has begun to examine promising dynamic risk factors for their ability to predict risk for future offending. The implicit assumption with dynamic risk factors is that risk is highest and most imminent when dynamic risk factors are present and active (e.g., violence risk is greater and more imminent when the individual is intoxicated; Klepfisz et al., 2016). Therefore, dynamic risk factors will hypothetically be better at predicting imminent and short-term risk for violent recidivism. To test this premise, some research has compared the short-term predictive validity of static and dynamic risk factors.

Validation studies of this sort typically describe a tool's predictive accuracy using the area under the curve (AUC) statistic, which is a commonly used and

recommended statistic as a measure of effect size for predictive recidivism (Rice & Harris, 2005). AUCs evaluate the trade-off between sensitivity and specificity of a tool. The benefit of AUC statistics is that they are not affected by the base rate of recidivism. AUC values can range between 0 (perfect negative prediction), through .5 (chance prediction), to 1 (perfect positive prediction). Rice and Harris (2005) have suggested that an AUC of .56 indicates a small effect size, .64 suggests a moderate effect, and .71 reflects a large effect size. As an analogy, Cohen (1969) suggests that a small effect is real but barely perceptible, such as the difference in average height between 15- and 16-year-old girls (about ½ an inch). A medium effect is likened to the easily visible difference in the average height of 14- and 18-year-old girls (about one inch), and a large effect is an obvious average difference in the heights of 13- and 18-year-old girls (about two inches; Cohen, 1969).

Chu et al. (2013) used the AUC statistical analysis to test the hypothesis that dynamic risk factors will be better at predicting imminent and short-term risk for violent recidivism. They compared several well-known risk tools that are based more on static risk factors (e.g., VRAG) to tools that feature dynamic risk factors (e.g., HCR-20). They examined and compared predictive validity at one-, three-, and six-month follow-ups. Their sample consisted of 66 male and female inpatients at a high-security forensic hospital. They hypothesized that dynamic risk measures would outperform static risk tools at the one- and three-month (short-term) follow-up periods. Their findings supported this hypothesis, indicating that historical risk tools were inferior at risk prediction over these short-term increments. That is, the dynamic risk scales had an average AUC of .74 and .75 at one- and three-months, respectively, and the static risk

scale had an AUC of .67 and .64 at one- and three-months, respectively (Chu et al., 2013).

In 2016, Finch et al. attempted to replicate Chu et al.'s (2013) findings with a different sample. In contrast to Chu et al.'s (2013) sample of high-security forensic inpatients, Finch et al. (2016) sampled 37 low-security forensic inpatients and 37 civil inpatients. They retrospectively coded files to compare the same (static or dynamic) risk instruments as the previous study using the same follow up periods of one-, three-, and six-months. Finch et al. (2016) replicated Chu et al.'s (2013) finding that dynamic risk measures were better than static risk measures at predicting aggression and violence in the short-term (i.e., one- and three-month follow-up periods), thus supporting the hypothesis that dynamic risk factors are stronger at predicting more imminent risk. Interestingly, a growing body of literature has suggested that dynamic risk factors may also add incremental predictive validity beyond static factors in medium- and long-term risk prediction, likely by tapping into stable dynamic risk factors (Beggs & Grace, 2010; Chu et al., 2013; Craissati & Beech, 2003; Desmarais et al., 2012; Finch et al., 2016).

The area of dynamic risk assessment is still a relatively new, yet promising field (Beggs & Grace, 2010). Research has clearly demonstrated that using evidence-based risk tools is far superior to subjective judgements of risk, and it is worthwhile to invest further research into improving the tools used for risk-related decision-making (Campbell, Webster, & Glass, 2009; Gendreau et al., 1996; Hanson & Bourgon, 2009; Monahan, 1981). This should include further investigation into the value of dynamic risk factors in the risk assessment process for both risk assessment in general, and IPV risk assessment specifically.

IPV Risk Assessment

As with general risk assessment, the use of first-generation, unstructured professional judgement is considered too unreliable and is, therefore, discouraged in IPV risk assessment. Instead, IPV risk assessment tools are a key piece of the risk assessment process, providing information to decision-makers to inform and guide the course of action they choose. There are a variety of second- and third-generation IPV-specific risk assessment tools available for different types of professionals working within the Canadian criminal justice system. These IPV-specific tools can be used to provide a common language, improve decision-making, and ensure proper resource allocation in cases of domestic violence. Some IPV risk tools are used to predict recidivism dichotomously, others are used to provide more detailed information about the nature, form, and severity of the risk for violence, and still others are designed to inform violence prevention and risk management (Kropp, 2008).

Second-Generation IPV Risk Assessment Tools

For many years, Harris et al.'s VRAG (1993) was the “go-to” second generation, actuarial risk assessment tool used for IPV, despite its intended use for general violence recidivism. The VRAG was developed by examining the predictive value of various correlates with violent recidivism, which resulted in a 12-item tool. Nine of the items are static and unchangeable (e.g., lived with biological parents until age 16, ever married, elementary school problems), and three of the items are stable dynamic (e.g., meets criteria for schizophrenia). The use of the VRAG for IPV risk assessment was justified by the fact that IPV offenders were included in the original sample that the VRAG was developed with, and the fact that preliminary research seemed to support such use

(Harris et al., 1993; Quinsey et al., 2006). That is, of the 500 violent offenders in the development sample, 88 had assaulted their intimate partner. These IPV offenders showed significantly lower-than-average VRAG scores and recidivated at a rate 20% lower than the other offenders of the sample (Hilton et al., 2001). Several other studies have also demonstrated the VRAG's ability to predict IPV recidivism (Hanson, Helmus, & Bourgon, 2007; Storey et al., 2011).

Despite the use of the VRAG by many police services, Hilton, Harris, and Rice (2010) describe receiving feedback directly from police officers that many of the best predictors on the VRAG required information not routinely available to them, such as early childhood problems, clinical assessment results, and juvenile delinquency records. In response to this limitation, they decided to develop an IPV actuarial risk tool specifically designed for frontline law enforcement that would be easy to administer, with the goal of helping to guide police decisions when responding to calls for IPV (Hilton, Harris, and Rice, 2010). Implementing the same statistical approach that was used to develop the VRAG, Hilton et al. (2004) examined police files to look for predictors of repeated IPV that police officers could easily score. These efforts ultimately led to the development of the Ontario Domestic Assault Risk Assessment (ODARA).

The ODARA is considered a second-generation tool because its 13 items were strictly chosen based on their statistical relevance, and its items are mostly static/historical in nature (Bonta & Andrews, 2007; Hilton et al., 2004; see Appendix A). Eight of the ODARA items are static in the sense that they have “happened” and that fact cannot change over time (i.e., ‘prior IPV,’ ‘prior other conviction,’ ‘prior sentence

of 30 days or more,' 'prior breach of court order,' 'victim and/or offender have > 1 child,' 'victim has child with previous partner,' 'previously violent outside of intimate relationships,' and '(previous or current) assault on victim while pregnant'). The 'substance use' and 'victim barriers to support' items can be static, dynamic, or both, based on the way the ODARA training manual describes it (Mental Health Centre Penetanguishene Research Department & Ontario Provincial Police Behavioural Sciences Section, 2005). For example, both item descriptions include consumption of alcohol during the index event (dynamic) and previous problematic substance use (static). There are three items that could be considered dynamic risk factors, including 'threats to harm or kill during the index,' 'confinement of the partner at index event,' and 'victim fears or is certain of future violence.' Unlike dynamic risk items in third-generation risk tools, the dynamic items of the ODARA were chosen based solely on their statistical significance, not on their theoretical or practical relevance to IPV behaviour.

The items on the ODARA are scored dichotomously as either present (1) or absent (0), and the sum of item scores provides a risk rating. Each score corresponds to a percentile rank in the referent population and an estimated probability of IPV recidivism (Hilton et al., 2004). In addition, risk scores correspond with one of seven categories of risk levels. At the extremes, a score of zero represents the least likelihood of recidivism, with 11% of Hilton et al.'s (2004) normative sample falling into this category, and a score of seven or higher represents the highest risk category, with only 7% of the sample falling into this group. Higher scores indicate a higher likelihood that an offender will recidivate more frequently, more severely, and sooner than those with lower scores.

Hilton et al. (2004) examined interrater reliability in this original cross-validation study using 24 cases, finding an intraclass correlations (ICC) of .90 for the index ODARA score and .91 for the recidivism ODARA score between two double-blind research assistants. In addition, two police officers (who were uninvolved in the research) demonstrated an ICC of .95 after coding 10 ODARAs independently (Hilton et al., 2004). Research findings have suggested that the ODARA is equal, if not superior, to other popular IPV risk tools in predictive accuracy. For example, Messing and Thaller (2013) published a meta-analysis that examined almost a decade worth of research on IPV risk assessment. They found the ODARA had the largest effect size of six commonly used IPV risk assessment tools, with an average AUC of .67 ($k = 5$).

Third-Generation IPV Risk Assessment Tools

Third-generation IPV assessment risk tools come with improved practical utility by emphasizing both static and dynamic risk items, which are selected based on both theoretical and general empirical support (Northcott, 2012). The increased emphasis on dynamic risk factors allows for treatment planning and risk management, working towards the goal of violence prevention. The Spousal Assault Risk Assessment Guide (SARA; Kropp et al., 1999) and the Brief Spousal Assault Form for the Assessment of Risk (B-SAFER; Kropp & Hart, 2004) are examples of commonly used third-generation IPV risk tools, which are both considered SPJ tools.

The SARA, a Canadian tool developed by the British Columbia Institute Against Family Violence to predict both IPV recidivism and lethality, is currently being used in 15 countries around the world (Helmus & Bourgon, 2011; Northcott, 2012). The SARA is comprised of 20 items related to psychosocial adjustment, general criminal history,

IPV history, and information about the current offence. Nine of the SARA's items are dynamic risk factors, thereby giving the SARA the ability to guide decision-making about treatment plans, interventions strategies, and required levels of supervision (Guo & Harstall, 2008). For example, one of the dynamic SARA items asks about 'recent employment problems,' which can be targeted in an intervention plan by helping the offender develop work skills and find gainful employment (Andrews & Bonta, 2003; Bonta & Andrews, 2007; Kropp et al., 1999). Other dynamic risk items include 'recent relationship problems,' 'recent problematic substance use,' 'recent suicidal/homicidal thoughts or intent,' 'recent psychotic or manic symptoms,' 'personality disorder with anger, impulsivity, or behavioural instability,' 'recent escalation in frequency or severity of assault,' 'extreme minimization of IPV,' and 'attitudes that condone violence.' The SARA requires the assessor to determine if each of the factors are present and if each is relevant to the case at hand. Information is gathered through an interview with the offender and the victim, the use of standardized assessments of emotional, physical, and problematic substance use, and collateral sources (e.g., informant reports, police records; Kropp et al., 1995). The assessor then considers all the data and makes a judgment of risk. A new version of the SARA with updated normative data has been recently published, which also requires the rater to make SPJ ratings on several potential outcomes (e.g., case prioritization, risk for severe violence; Kropp & Hart, 2015).

Interpreting research on the effectiveness and accuracy of the SARA is complicated by variations in the methodologies used. Although the tool was designed to be used as an SPJ assessment informed by many sources (e.g., interviews, file review, standardized test data), many studies examine its predictive validity using file coded,

actuarial risk scores (i.e., the summed count of individual item ratings). For example, in a review of 11 studies on the SARA from 2000 to 2008, Helmus and Bourgon (2011) reported the SARA's AUCs separately for different methodologies to make sense of the results. Five of these studies in their review examined the SARA summary risk ratings (i.e., the intended SPJ approach), finding an average weighted AUC of .67; however, the meaning of this average is questionable given the significant variability in effect sizes across studies, ranging from .56 to .77. The authors noted that this large range in effect sizes likely reflects true variation, given the SPJ approach is impacted by the subjectivity of the professionals making the judgements. Another four studies in this review examined the predictive validity of the SARA as an actuarial tool (i.e., item scores were totaled mechanically). The average weighted AUC of these studies was .63. Finally, four additional studies compared the SPJ and actuarial approaches to the SARA within the same sample, finding three studies favoured the SPJ approach and one favoured the actuarial approach. Of note, the follow-up periods of these 11 studies also ranged from 3 to 60 months, but the authors did not examine the impact of length of follow-up time on the predictive ability of the SPJ and actuarial SARA. A later meta-analysis conducted by Messing and Thaller (2013) examined the predictive validity of the SARA across six studies, finding a moderate weighted average AUC of .63, with AUCs ranging from .59 to .65. They also suggested using caution in interpreting the average AUC given the variability in methodologies across studies.

The strengths of the SARA include the fact that it has the most independent validation studies to date, and it was developed based on diverse empirical research and, therefore, has high generalizability across sub-populations. Another strength of the

SARA is that its inclusion of both static and dynamic risk factors makes it sensitive to change in offender risk and useful for treatment planning (Helmus & Bourgon, 2011). Limitations of the SARA include the fact that it is quite lengthy and because of its subjective nature, the interrater reliability of its items varies greatly and is moderate (with ICCs ranging from .45 to .86; Kropp & Hart, 2000). Overall, the research on the SARA indicates that it is promising, but more research is needed to understand its value as a SPJ tool versus as an actuarial tool.

The SARA has been condensed into a shorter tool based on the results of exploratory and confirmatory factor analyses, resulting in the Brief Spousal Assault Form for the Evaluation of Risk (B-SAFER; Kropp & Hart, 2004). The B-SAFER is comprised of 10 items, with 4 static items (e.g., ‘previous serious violence’ and ‘previous violations of court orders’) and 6 dynamic items (including ‘attitudes condoning IPV,’ ‘relationship problems,’ ‘escalation of violence,’ ‘employment or financial problems,’ ‘problematic substance use’ and ‘mental disorder’). Each item is coded as present, possibly/partially present, and absent. Like the SARA, the B-SAFER is a SPJ tool. Therefore, the final step in coding the B-SAFER is for the evaluator to make a judgment of imminent risk, long-term risk, and lethal violence risk (with each of these judgements coded as high-, moderate-, or low-risk; Kropp & Hart, 2004).

As part of their pilot study, Kropp and Hart (2004) found that police who used the B-SAFER recommended more risk management strategies in cases that were higher risk, a promising finding for the utility of the tool. Au et al. (2008) conducted a preliminary validation study in Hong Kong, finding that the B-SAFER could correctly identify 90% of the IPV offenders and 100% of the controls in a sample of 43 male IPV

offenders and 46 controls. Storey and Strand (2013) examined 53 Swedish cases of female-perpetrated IPV and the corresponding B-SAFER risk judgments as scored by police officers. They compared their findings to those of another study using the sample that examined male IPV perpetrators (i.e., Belfrage & Strand, 2008). They found that officers assigned higher summary risk ratings to male IPV perpetrators, even when the mechanically calculated B-SAFER totals were the same for men and women. They also found the average number of risk strategies recommended by police officers only increased with risk for male perpetrators, not for female perpetrators. This suggests that the subjective risk judgments were biased by police officers' ideas regarding gender and risk (Storey & Strand, 2013).

The benefits of the B-SAFER over the SARA include the fact that it takes less time and fewer resources to administer, it has less technical jargon, and it can be used by non-mental health professionals (Kropp & Hart, 2004). It does not seem to compromise accuracy in its briefer format compared to the SARA, but more research on the B-SAFER's validity and reliability is needed. Like the SARA, the B-SAFER may be limited by the subjectivity of SPJ tools, which can lead to lower interrater reliability (Guo & Harstall, 2008). For example, Thijssen and de Ruiter (2011) examined the interrater reliability of the B-SAFER in a Dutch sample, finding the intraclass correlations were only moderate (the average ICC was .57), with the mental health item showing particularly poor interrater reliability (ICC = .21).

Overall, it is clear that the VRAG, ODARA, SARA, and B-SAFER all have strengths and limitations as IPV risk tools. In terms of police use of IPV risk assessment tools, a shared limitation of the VRAG, SARA, and B-SAFER is that they necessitate

access to various sources of information not routinely available to police officers. For example, the VRAG and the SARA both require information about the suspect's childhood. The SARA and the B-SAFER both require clinical information about the suspect's mental health or personality diagnoses, which is generally not practical for frontline police use due to limited time, resources, and clinical training. One of the strengths of the ODARA over these other tools is that because it was developed specifically for frontline police officers, it is relatively quick and easy for police to score (Guo & Harstall, 2008; Nicholls et al., 2013). It is free from technical jargon, does not require clinical training or qualifications, and can be scored using information that is accessible to police (e.g., victim interview, suspect interview, police records database; Hilton, Harris, and Rice, 2010). As an actuarial tool, it also provides the structure and objectivity that many police officers prefer (Moser, 2012).

The ODARA, however, is not without limitations. That is, the ODARA's emphasis on static risk limits its ability to identify risk factors to be targeted for the prevention of future violence. It also means that it is not particularly sensitive to changes in risk (Moser, 2012). This means that an offender's risk score is not likely to ever decrease after their baseline ODARA score has been established (i.e., risk status), limiting its utility for detecting escalations in risk that may lead to preventable violent outcomes. Similarly, it is limited in its utility for police to examine the effectiveness of their responses and risk management decisions.

Police Use of IPV Risk Assessment Instruments

The primary goal of IPV risk assessment use by police is to prevent future IPV (Hart, 2010). According to the RNR model, this can be accomplished by using risk

assessment instruments to derive an estimate of a perpetrator's risk for future IPV, identify dynamic risk factors relevant to the situation, and use risk management strategies to address the perpetrator's criminogenic needs at a level of intensity appropriate for their level of risk. For example, if an IPV perpetrator uses substances while assaulting their partner, then a police officer could issue a condition of release that they abstain from using substances. Alternatively, if an officer observes an escalation in IPV severity or frequency over time, then they might seek to hold the perpetrator for court to protect the victim from imminent harm. If the victim expressed fear or certainty that their partner will continue to abuse them in the future, then the officer could refer them to victim services, order the perpetrator to have no contact with the victim, or safely transport the victim to a shelter or safe house.

Although the use of risk assessment instruments is quickly becoming a more regular part of policing IPV, there is very limited research on their ability to guide police decision-making in frontline contexts and to prevent recidivism. In one study, researchers examined almost 700 cases of IPV in which Swedish police officers used the B-SAFER on the frontline (Belfrage & Strand, 2008). They found that the intensity of the police response was strongly associated with the number of perpetrator risk factors present. Similarly, Trulijo and Ross (2008) found that police based their decisions about the intensity of their chosen interventions on the presence of a few specific risk factors (i.e., escalation and victim fear). An important next step in the research on police use of IPV risk assessment instruments is to examine the potential mediation effects of police risk management strategies on the relationship between assessed risk level and recidivism.

Belfrage et al. (2011) conducted a prospective study of police response to IPV in Sweden by auditing all IPV cases in three counties in the year 2000 and for 18 months thereafter. Police in these counties were trained on the SARA and used it to assess risk and inform risk management plans. Using a final sample of 429 IPV cases, they found that the police-scored SARA totals and summary risk ratings were significantly correlated with the number of risk management recommendations made by police (both $r = .40, p < .001$), such that officers made more intervention recommendations for higher risk cases. They also examined the mediation effect of risk management on the association between risk assessment and recidivism. In line with the RNR model, they found high-risk perpetrators with many risk management recommendations were *less* likely to recidivate than high-risk perpetrators with few risk management recommendations. Conversely, low-risk offenders with many risk management recommendations were *more* likely to recidivate than low-risk offenders with fewer risk management recommendations.

In a similar study, Storey et al. (2014) examined 249 IPV cases in which Swedish police used the B-SAFER to assess risk and develop risk management plans. They found that total scores and overall risk ratings predicted recidivism with a relatively large effect size (AUCs of .70 and .69, respectively). In addition, they found risk management mediated recidivism, such that more risk management recommendations corresponded with decreased recidivism in high-risk offenders. Consistent with Belfrage et al.'s (2011) findings, they also found that more risk management recommendations increased recidivism in low-risk offenders, a finding that supports Andrews and Bonta's (2003) RNR framework for offender rehabilitation. As

with any research, studies such as those which were done in Sweden by Belfrage et al. (2011) and Storey et al., (2014) need to be replicated with different samples to ensure the findings are reliable across populations and contexts. Sweden and Canada, for example, have very different cultures. Sweden is a world leader in gender equality in politics, big business, and general workplaces. However, Sweden has an unusual situation regarding IPV. There is a major deficit in services available for victims and they generally have a “don’t ask, don’t tell” attitude towards IPV (Pratt-Eriksson et al., 2014). Replicating research from Sweden on Canadian samples can give us more confidence that the findings are relevant to Canadians.

Current Study

The current research had three primary goals. First, it aimed to evaluate police use (i.e., frequency of use, scoring, predictive validity, application to risk management) of one of the most commonly used IPV risk tools in Canada: the Ontario Domestic Assault Risk Assessment (ODARA; Hilton et al., 2004; Millar, 2009). The ODARA is being systematically used by police forces across Canada, including in Nova Scotia, New Brunswick, Ontario, and Saskatchewan (Millar, 2009). The second goal of the current study was to develop a dynamic risk tool that police officers could use as a supplement to the ODARA. The third research goal was to do a preliminary examination of said dynamic risk supplement to the ODARA to determine if it would improve short-term risk prediction and sensitivity to change in risk over time.

Leveraging the benefit of the widespread implementation of the ODARA in the province of New Brunswick, the current study used a New Brunswick sample to engage in the planned research goals. The process of provincial adoption of the ODARA onset

when the Domestic Violence Death Review Committee of New Brunswick made a series of recommendations in 2014 to the Government of New Brunswick regarding proactive and reactive responses of the criminal justice system about the issue of domestic violence (Public Safety Canada, 2017). They issued a recommendation that all law enforcement officers in the province undergo mandatory training on IPV and IPV risk assessment, as well as a recommendation that IPV risk assessment by police officers be considered mandatory. The New Brunswick Department of Public Safety responded to these recommendations by encouraging the New Brunswick Policing Standards to require all police services in the province to follow the “Woman Victims of Abuse Protocols” (Office of the Chief Coroner, 2015).

Technically, the “Woman Victims of Abuse Protocols” speak to male-perpetrated abuse against women; however, the principles, definitions, and protocols within it can be applied to both male and female victims of IPV and those in same-sex relationships. The definition of IPV set out in these protocols includes several forms of chargeable and nonchargeable behaviours within the categories of psychological, emotional, verbal, physical, sexual, financial, and spiritual abuse (Women’s Equality Branch, 2014). In terms of its application to law enforcement, the definition is narrowed to chargeable offenses within those categories. In Canada, there are no criminal code offenses specific to family violence or IPV, but most IPV behaviours fall under other offense categories, such as assault, threats, harassment, extortion, and so on (Department of Justice, 2021). The “Woman Victims of Abuse Protocols” specify that when police receive a complaint of a chargeable offense that occurred between current or former intimate partners, they must conduct an investigation, complete and document a

validated risk assessment tool, and take the suspect's level of risk into consideration when making any decisions (Women's Equality Branch, 2014).

Following the development of these protocols, the New Brunswick Association of Chiefs of Police agreed to the province-wide adoption of the ODARA by frontline police officers when responding to situations of IPV (Office of the Chief Coroner, 2015). By March 2015, all police officers in New Brunswick were trained on the ODARA, and it became protocol to use the ODARA for every call of IPV in which a chargeable offense took place, regardless of suspect gender or sexual orientation (A. Totten, personal communication, December 13, 2017).

Research Goal One

Several factors were explored in the evaluation of the use of the ODARA by police officers in the current study, such as how often officers completed the ODARA in the cases used in the current sample (i.e., for what proportion of IPV calls was it completed), and how it was scored. This evaluation also sought to examine if certain factors would influence when/for whom the ODARA was completed. The first three variables examined were suspect gender, suspect and victim relationship status, and living arrangement. These were chosen because the ODARA was originally developed to be used only with male suspects and their current partners with whom they reside (Mental Health Centre Penetanguishene Research Department & Ontario Provincial Police Behavioural Sciences Section, 2005). Many officers would have originally been trained to use the ODARA within these limits; however, in 2015, the sampled police service decided to use the ODARA with female suspects, partners who had broken up, and partners who were not living together (based on emerging research support for such

use; A. Totten, personal communication, December 13, 2017). The remaining four variables examined were victim cooperation with the investigation, suspect use of a weapon during the index event, suspect use of substances during the index event, and whether there was a visible victim injury. These variables were chosen because previous (published and unpublished) research using a comparable sample of police officers from New Brunswick have found these four variables to have notable influence over police decision-making when responding to calls of IPV (Ballucci et al., 2020; Moser, 2012).

The current study also explored how police use ODARA risk scores to guide their decision-making by examining if the police-calculated risk scores predicted the intensity of their intervention. Based on findings of previous research, it was hypothesized that higher police-calculated risk scores would be associated with a greater volume of risk management strategies and higher intensity intervention (Belfrage & Strand, 2008; Trulijo & Ross, 2008). Finally, the current study explored whether risk management mediated the relationship between police-calculated risk scores and recidivism. Based on the RNR model and previous research supporting it, it was hypothesized that risk management would mediate the association between risk assessment score and recidivism (Belfrage et al., 2011; Storey et al., 2014).

Research Goal Two

The second goal of the current study was to develop a supplement to the ODARA that had a greater emphasis on dynamic risk for IPV recidivism, but which retained the ODARA's intent to be scored quickly with information that is routinely accessible to frontline police officers. To identify the most appropriate factors to include in the supplementary tool, all dynamic risk factors that previous research has

consistently identified as either significantly correlated with or predictive of IPV recidivism was coded from police files when information was available (see review of dynamic risk factors for IPV in introduction above). These coded potential dynamic risk factors included: escalation of IPV, antisocial attitudes, relationship problems, employment/financial problems, problematic substance use (i.e., a purely dynamic alternative to the version of this item in the ODARA), suicidality/self-harm (representing an extreme emotional state, which can be operationalized and identified by police officers more easily than mental health concerns or diagnoses; Blosnich & Bossarte, 2012; Coker, Davis, et al., 2002; Flynn et al., 2016; Hart, 1990; Hilton et al., 2001), victim fear (an existing ODARA item), and excessive jealousy/control. These variables are operationalized and discussed in more detail in the “Method” section below.

The potential dynamic risk factors were evaluated based on Douglas and Skeem’s (2005) first two criteria to provide an indication of whether they are truly criminogenic needs (i.e., dynamic risk factors are correlated with IPV and changeable over time). The current study did not set out to examine Douglas and Skeem’s third criterion (i.e., changes in the dynamic risk factors must influence recidivism), which is a noted limitation of this dissertation. Each suspect would need to have had a minimum of three police contacts during the 12-month follow-up period to evaluate if change in dynamic risk between ‘contact one’ and ‘contact two’ can predict the likelihood of IPV recidivism at ‘contact three’ (Andrews & Bonta, 2003). The logistical constraints of this dissertation project would not have allowed for a large enough sample of people who had at least three IPV-related police contacts. In addition, because only police file data

was accessed for the current study, the analyses were limited only to those who recidivated. Examining this third criterion is a complex empirical endeavour that would be better addressed in a large-scale prospective study that has access to data on both recidivists and non-recidivists (e.g., a probation sample), with a longer follow-up period. However, the current study explored and examined many other questions that offer unique contributions to the field of IPV research and lays the groundwork on which to build future research.

Research Goal Three

The third goal of the current study was to compare the dynamic risk supplemental tool (as a third-generation tool) to the ODARA (as a second-generation tool). The first objective was to examine the ability of the dynamic risk supplemental tool to reliably measure change in an individual's risk state over time, and it was hypothesized that it would provide incremental validity over the ODARA's ability to do the same. The second objective was to compare the predictive validity of the dynamic supplement and the ODARA over the 12-month follow-up period generally, as well at multiple time points that were chosen to replicate the timeframes used in previous research on dynamic risk factors (Chu et al., 2013; Finch et al., 2016). The follow-up times to be used included less than one month (imminent risk), one to three months (short-term risk), three to six months (medium-term risk), and up to 12 months (medium-to long-term risk).

The current study also aimed to compare the ODARA with the dynamic supplement at predicting how quickly (in days) suspects engage in IPV recidivism (i.e., time to recidivism). Given that risk is highest when dynamic risk factors are present

(Klep̄isz et al., 2016), they should, by definition, be stronger at predicting imminent and short-term risk recidivism. Research on general violence risk assessment instruments has supported this hypothesis using both forensic and community samples (Chu et al., 2013; Finch et al., 2016; McNeil et al., 2003). Therefore, it was hypothesized that the dynamic supplement would (a) show improved accuracy in predicting imminent recidivism (<1 month), short-term recidivism (<3 months), and time to recidivism (measured in days), and (b) not outperform the ODARA at the medium- (6 months) and long-term (12 month) follow-up periods.

Notably, although previous research has established that the ODARA is useful for predicting the likelihood that an IPV offender will recidivate at some time in the future, the ability of the ODARA to predict imminent or short-term risk specifically has not yet been examined. Existing research on the ODARA has an average follow-up period of five years (Hegel et al., 2021; Hilton, Harris, & Rice, 2010). Likewise, no published research has examined the ODARA's ability to detect changes in risk over time. These gaps in the previous research reflect the historical context of risk assessment. In the past, risk assessment tools were designed for risk classification purposes (i.e., if the person will eventually recidivate). Interest in (dynamic) risk state and the changeability of risk is a relatively new focus; therefore, this has not been a priority in previous research. However, risk state and imminent/short-term IPV risk prediction could be extremely valuable for frontline officers making risk management decisions and safety planning for victims.

Method

Sample

The sample for the current study was drawn from the IPV-related reports generated by a municipal police service in New Brunswick, Canada. The mandatory use of the ODARA for IPV calls was implemented by the New Brunswick Chiefs of Police on January 1, 2015, and all members of their partnering police forces (including the police force used in the current sample) were trained on the ODARA by March 1, 2015 (Office of the Chief Coroner, 2015). Given the implementation of this new protocol and the recency of the IPV and ODARA training, the objective was to draw index files from March 1, 2015 to February 29, 2016. The initial intention behind the selection of this period was that it would, in theory, show a higher rate of officer compliance to the ODARA protocol and accuracy in scoring given its closeness to the date of training, which would, in turn, generate useful data for analyses involving examination of police use of the ODARA.

To develop the sample, the crime analyst for the partnering police service randomly pulled relevant police files for the researchers (i.e., files identified by the responding officer(s) as involving intimate partners over the age of 18). Each selected case was reviewed by the researchers to determine if it met seven inclusionary criteria (described below). If a file did not meet these criteria, then it was discarded from the sample and a new case was randomly selected to replace it. These cases represent IPV index events (i.e., baseline events from which initial measurements were taken). Although the “index” event was the first event coded for the purpose of the current study, it was not necessarily the first IPV offence for that individual. Thus, when the information was available, it also was noted if the suspect had a documented history of IPV prior to the index event.

The intention of current study was to use this random selection process until 309 index cases met the inclusion criteria, in order to ensure sufficient power for planned analyses. Unfortunately, there were not enough qualifying files within the intended timeframe (i.e., March 1, 2015 to February 29, 2016), so the timeframe was extended forward until 12 months before the date at the time of the data collection (i.e., in order to allow for a full 12-month follow-up period for each suspect). However, the data collected for the current dissertation were part of a larger research project that involved the study of the role of gender in IPV risk prediction. As it turned out, there were still not enough qualifying files within this broader timeframe that had female suspects to meet the minimum sample size requirements for this secondary project. As such, some quota sampling was used to ensure a minimum of 100 female participants were in the sample. This quota required files to be drawn from dates prior to March 2015. Therefore, all analyses that examined police use of the ODARA (i.e., research question one) were restricted to only those cases for which the index file was from March 2015 or later (i.e., post-ODARA training, referred to as “post-2015 files”), which used random sampling. All other analyses (i.e., research questions two and three) relied on the ODARAs completed by the researchers using file data drawn from the broader time frame (i.e., the full sample), 62% of which was derived from random sampling and 38% randomized quota sampling (randomly drawn files with female suspects). In the end, the files for this research project were drawn from June 2011 to January 2017.

The seven inclusionary criteria used to ensure that each file was appropriate for use were as follows: 1) only English language files were included in the study due to unilingual limitations of the researchers (although this ended up being irrelevant as all

the files were in English); 2) only files in which the perpetrator and victim roles could be unambiguously identified were included; 3) both parties were over the age of 18 at the time of the index event; 4) the victim and perpetrator referred to in the file were either currently or previously in an intimate relationship with each other; 5) police intervention was a direct result of an issue between the two parties of the intimate relationship; 6) the responding officer must have indicated in the file that there was some form of domestic dispute between the victim and perpetrator, with the perpetrator's aggression intentionally directed at the victim; and 7) the domestic dispute must have met the operational definition of IPV used in the current study.

How to operationally define IPV is not unanimously agreed upon. For instance, the definition of IPV in the ODARA development manual is an event in which a man has assaulted his female partner and/or her children (Mental Health Centre Penetanguishene Research Department & Ontario Provincial Police Behavioural Sciences Section, 2005). Others have argued that IPV transcends gender and sexual orientation and, therefore, should be defined more broadly as actual or threatened abuse by a current or former spouse, common-law, or dating partner, regardless of the genders of those involved (Dawson & Hotton, 2014; Whitaker & Lutzker, 2009). A case can certainly be made for why the developers of the ODARA chose to use a gendered definition of IPV, given that it was normed on convicted male perpetrators of IPV. In addition, there is typically a massive gender bias in police-reported IPV, such that 80% of cases involve male suspects and female complainants (Statistics Canada, 2016), and subsequently there are disproportionately more males than females convicted of IPV. Thus, the ODARA definition of IPV as male-on-female violence would have been the

most common scenario and, therefore, a convenient sample to access. However, research shows that the rates of IPV are similar regardless of gender when self-report measures are used (Statistics Canada, 2016). Although most validation research on the ODARA has focused on male perpetrators, preliminary research suggests that the ODARA may be able to accurately predict recidivism in female perpetrators as well (Hilton et al., 2014; Moser, 2012; Olver & Jung, 2017). Thus, there is justification to include female perpetrators in validity studies of the ODARA in an effort to replicate these findings.

Notably, the ODARA's definition of IPV is also limited to physical violence (although they included one type of sexual abuse: use of force to coerce unwanted sexual contact), and most research on IPV to date has focused on physical violence (Cale et al., 2017; Outlaw, 2009). However, several scholars have argued that the definition of IPV should also include a broader spectrum of sexual abuse, psychological/emotional abuse, stalking, and/or threats of violence (Whitaker & Lutzker, 2009). Interviews with IPV victims indicate that their experience of these other types of often IPV are far more impactful on their daily lives and functioning (Strauchler et al., 2004). Other research suggests that the damage caused by non-physical forms of IPV can be equal to, or greater than, that caused by physical violence (Coker, Davis, et al., 2002; Mechanic et al., 1996; O'Leary, 1999; Rogers & Follingstad, 2014). In addition, non-physical IPV is far more common than physical IPV (Outlaw, 2009). Preliminary research has demonstrated the ODARA's ability to discriminate between IPV recidivists and non-recidivists regardless of whether the IPV was physical or non-physical (Moser, 2012).

In an effort to replicate Moser's (2012) findings showing that the ODARA is able to accurately distinguish recidivists from non-recidivists regardless of suspect

gender, victim gender, or type of violence committed (i.e., physical versus nonphysical), an expanded definition of IPV was used in the current research. Specifically, the operational definition of IPV used in the current research was “an event in which a current or former spouse, common-law partner, or dating partner abuses, threatens to abuse, or attempts to abuse their intimate partner by use of physical abuse, sexual abuse, psychological/emotional abuse, stalking/harassment, or financial abuse, regardless of the gender of either persons involved.” This definition is consistent with the definition that was developed and used by the federal/municipal/regional police forces across New Brunswick, the Domestic Violence Court in Moncton, the prosecution services of the Office of the Attorney General, and New Brunswick’s Crime Prevention and Reduction Strategy at the time from which these data were collected (New Brunswick Multicultural Council, 2018; Province of New Brunswick, 2012). In addition, this is the definition of IPV that officers from the police force in the current study were trained to use in their application of the ODARA, with an implied emphasis on chargeable behaviours (i.e., the training completed by March 2015; A. Totten, personal communication, December 13, 2017).

Although the role of gender was not an explicit focus of the current dissertation, exploratory analyses were conducted prior to all subsequent analyses to examine whether there were significant differences in the results for male and female suspects of IPV given the use of quota sampling. The results are reported with the aggregate data (collapsed across gender) whenever the patterns of results were very similar for both male and female suspects. The final sample included 209 males and 100 females. The mean age of the suspects was 33.5 years ($SD = 10.9$). The full sample was comprised of

85.4% Caucasian suspects ($n = 264$), 5.5% Black suspects ($n = 17$), 2.9% Indigenous suspects ($n = 9$), 2.6% suspects of another ethnicity/race ($n = 8$), and 3.6% suspects of an unknown racial background ($n = 11$).

Measures

Ontario Domestic Assault Risk Assessment

The Ontario Domestic Assault Risk Assessment (ODARA; Hilton et al., 2004; see Appendix A) was the subject of the current evaluation. Whenever available, police-scored ODARAs were extracted from the case files for the index and recidivism events. The researchers also scored an ODARA for each of suspects in the sample, regardless of whether there was a police-scored ODARA in the file. Many Canadian studies have shown the ODARA can be reliably scored from archival data (Hilton, Harris, & Rice, 2010; Jung & Buro, 2017; Moser, 2012). To avoid bias in the data, the researcher-scored ODARAs were coded blindly (i.e., prior to reading the police-rated ODARA scores and prior to collecting the outcome data for each suspect's file).

The scoring protocol developed by Hilton et al. (2004) was used (i.e., 0 = item not present, 1 = item present, or ? for missing information) to score the ODARA. Item scores were summed together to calculate the total risk score, ranging from 0 to 13. Total scores of 0 to 2 indicate the lowest risk for recidivism, with only approximately 22% recidivating within 5 years. Scores of 3-6 represent a moderate risk for recidivism, with approximately 39% recidivating within 5 years. Scores above 7 represent the highest risk category, with approximately 75% of people in this group reoffending within 5 years (Mental Health Centre Penetanguishene Research Department & Ontario Provincial Police Behavioural Sciences Section, 2005). If five or fewer items could not

be scored, then the total score was prorated using a table of adjusted scores provided in the scoring guide (Mental Health Centre Penetanguishene Research Department & Ontario Provincial Police Behavioural Sciences Section, 2005). If greater than five items could not be scored, then the file was excluded from the study and another random file was drawn to replace it.

A notable limitation of the ODARA, as with many other risk assessment instruments currently used in the Canadian criminal justice system, is a lack of literature examining its predictive validity with BIPOC populations (Hegel et al., 2022; Olver, 2016; Shepherd & Lewis-Fernandez, 2016). This matter was illustrated in the case of *Ewert v. Canada* (2015), when a man of Métis heritage sued Correctional Service Canada for using risk assessment tools in his case that were not psychometrically validated for use with Indigenous people. Although the case originally resolved in Mr. Ewert's favour, after an appeal process, the Supreme Court of Canada ultimately ruled that the use of such measures was not unconstitutional, yet that it was imperative for risk assessment tools to be psychometrically validated for use with Indigenous people (*Ewert v. Canada*, 2018). The ODARA has very little validation data for Indigenous people to date; however, initial findings suggest it may have significantly lower predictive accuracy for Indigenous populations compared to White populations, although it still appears to be more accurate than human judgement alone (Buchanan, 2009; Hegel et al., 2022). Hegel et al. (2022) found that controlling for suspect age increased the predictive validity of the ODARA when used with Indigenous suspects, suggesting that the ODARA's heavy emphasis on static risk factors may be the source of the lower predictive validity with Indigenous suspects. This makes sense given that Indigenous

people in Canada are more likely than non-Indigenous people to have a criminal record due to the over-surveillance of Indigenous communities (Comack, 2012). Although there is a need for additional validation research on ODARA with Indigenous populations, unfortunately this need was not addressed in the current dissertation given the homogeneity of the sample obtained (i.e., there was an insufficient subsample of Indigenous suspects).

Coding Guide

A comprehensive coding guide was developed for the current study to capture the remaining variables of interest from police file data, loosely based on the coding guide used by Moser (2012; see Appendix B for full coding guide). Research goal one of the current study involved examining if any situational or demographic variables influenced whether police officers administered the ODARA. Thus, the first part of the coding guide recorded this information. The coded situational and demographic variables included: incident details (e.g., date and time of incident, if it was an index or recidivism event, what type of offence was committed, use of a weapon); victim demographic information (e.g., gender, age, employment status, cooperation with police investigation); information about the relationship between the victim and suspect (e.g., relationship status, living arrangement); and suspect demographic information and characteristics (e.g., age, gender, details about previous IPV).

The type of abuse perpetrated was also coded using a checklist of many possible abuse types that was developed for the current dissertation. From this checklist, five composite variables were created to track the general types of abuses that were perpetrated for each coded IPV incident: physical abuse, emotional/psychological abuse,

property abuse, stalking/harassment, and financial abuse. Although psychological abuse and stalking/harassment can overlap in many ways, for the purpose of the current research, emotional/psychological was defined as actions that were intended to degrade, humiliate, insult, control, or instill fear in the victim (e.g., threats of violence, name calling, screaming at the victim), and stalking/harassment was defined as actions that involved following, watching, monitoring, or initiating unwanted contact with the victim (e.g., phone calls, texts, showing up places the victim is, sneaking into their house, leaving notes). Physical abuse included any physically aggressive action that was intended to scare, hurt, kill, or physically restrain/control the victim (e.g., confinement, choking, hair pulling, rape, shooting, slapping, spitting, stabbing). Property abuse included any action intended to destroy or damage property of the victim in an effort to control, instill fear in, or otherwise negatively affect the victim (e.g., forced entry, breaking the victim's personal items, vandalism). Finally, financial abuse involved any action to control or harm the victim through financial means (e.g., stealing from victim, fraudulent behaviour in the victim's name, parasitic lifestyle).

Evaluating the police use of the ODARA also involved examining whether their produced risk score would predict risk management strategies employed by the responding officers. Therefore, the coding guide recorded the actions of police during and after responding to an IPV call. Police action items included (a) general criminal justice responses (e.g., completed the ODARA, looked up suspect criminal record, arrested suspect, conducted investigation, recommended charges to Crown, reported incident to Department of Social Development if children witnessed IPV), (b) victim-directed actions (e.g., referral to shelter, gave victim information about family court

procedures, developed safety plan), (c) suspect-directed actions (e.g., referred to community services, removed from scene, and issued unofficial warning), and (d) other general actions (e.g., increased police presence in area where IPV took place, visited residence to check in).

Finally, in service of the second goal of the current research (i.e., to develop a dynamic supplement to the ODARA), the coding guide included the eight potential dynamic risk factors identified through the literature review as relevant to IPV risk. For each of these items, the rater coded whether there was no evidence (0), partial or inconclusive evidence (1), or clear evidence (2) of the dynamic factor during or closely preceding the index event. Ordinal scoring is commonly used with dynamic risk assessments because of the increased flexibility it gives raters to capture fluctuations in severity and the fact that partial scores are often more realistic than dichotomous classifications. For example, the SARA ratings are also based on a 3-point scale (0 = absent, 1 = possibly or partially present, 2 = present; Kropp & Hart, 2000). The B-SAFER also uses ordinal scoring with three response options (Y = the factor was present, ? = possibly or partially present, N = the factor was absent; Kropp & Hart, 2004). If there was no information available in the file to score the item, it was coded as 'no evidence' (0). That is, the rating of zero was used to indicate both if an item was not observed or if it was not recorded in the file, as there was no way to distinguish this through retrospective file review (elaborated upon further in the "Discussion" section).

The first dynamic item coded was 'recent escalation in IPV,' which was defined as a pattern of abusive behaviour becoming more severe or frequent in the month prior to the current IPV situation. Unfortunately, attempting to capture increased frequency in

IPV proved to be difficult given the fact that this information was rarely recorded in police file notes and only incident reports were used to code information (i.e., not the suspect's full record); thus, interrater reliability for this item was poor. As such, this item was removed from further consideration in the current study (considered in more detail in the "Discussion" section). The second item was 'relationship problems,' examples of which included current or recent separation/divorce, frequent and serious conflict, one partner wanting to end the relationship, or repeated infidelity. The next dynamic risk factor was 'excessive jealousy or control of victim,' which included current or recent behaviours such as obsessive checking in on victim, becoming angry or hostile in response to the victim's associations with other people, accusing the victim of being unfaithful without reasonable grounds for the accusation, attempting to control most or all of the victim's daily activities, using threats/blackmail to make the victim compliant, or isolating the victim from friends and family.

The fourth dynamic risk factor in the coding guide was 'employment or financial problems,' which included being recently laid off, current or recent unstable employment/income, or use of community resources to meet basic needs (e.g., housing, food). The fifth item in this section was 'self-harming/suicidal behaviours.' Given that the intent of the ODARA (and the dynamic supplement) is to be scorable without clinical training and to be free of technical jargon (Hilton et al., 2004), this item was chosen to represent an operationalization of a mental health crisis or extreme emotional state. Other researchers and risk assessment developers have used this item conceptualization with success, as these behaviours have been shown to increase the risk for IPV in general and greater severity of victim injury (Kropp et al., 1999; Kropp &

Hart, 2004). Examples of suicidality or self-harm in the coding guide included current or recent cutting, head banging, skin burning, threats of suicide or self-harm, or suicidal gestures or attempts.

The sixth dynamic risk factor in the coding guide was ‘demonstrates attitudes that condone violence,’ which included current or recent violations of court orders, complete denial of responsibility, opposed to police presence/involvement³, minimization of significance or impact of IPV, or generally talking favourably about criminal behaviours (hereafter shortened as ‘antisocial attitudes’). The seventh dynamic risk factor in the coding guide was ‘substance use during or leading up to the current IPV event,’ which included the use of any alcohol, illegal drug, or misuse of prescription drug (for the purposes of an intoxicating effect) during or immediately preceding the IPV.

The eighth dynamic risk factor, ‘victim fear,’ was adopted from the ODARA. The ODARA defines victim fear as the victim explicitly making a statement that indicates concern, fear, worry, or certainty about future IPV (Hilton et al., 2010). It does not include inferences of victim fear made by the rater based on the victim’s reaction to

³ As mentioned in the previous footnote, demonstrating anti-authority attitudes and behaviours has long been considered one possible indicator of antisocial or procriminal attitudes (Andrews, 2015; Andrews & Bonta, 2006; Bonta & Andrews, 2007; Cooke, 1996; Mills et al., 2004). As such, being “opposed to police presence/involvement” was included in the operationalization of this variable in the current dissertation to be consistent with previous research and because it would be a factor police officers could easily assess in a frontline context. Various other indicators were also captured in the operationalization of the ‘antisocial attitudes’ variable that did not involve anti-police attitudes (such as violation of court orders, rationalizing or justifying IPV, victim blaming, threatening future crimes to meet current needs). It was not until long after the data collection was completed that I recognized that although being “opposed to police presence/involvement” *could* be an indicator of antisocial attitudes for some individuals, for others it could be the result of distrust or disdain towards the police due to a longstanding and ongoing history of racialized police practices (Ibrahim; 2020; Statistics Canada, 2020b). I was not able to find any previous research that has examined the “anti-authority” component of “antisocial attitudes” by suspect race, and unfortunately the data for the current dissertation were not coded in such a way that allowed for such an analysis. However, I was able to compare Caucasian and BIPOC suspects on the variable as a whole. The results of these analyses are presented in footnotes in the “Results” section below.

the IPV or protective actions they take afterwards. The benefit of excluding any inferences of fear by the rater is that it reduces subjectivity and “mind-reading” on the part of the police officer. However, sometimes a victim is unwilling or unable to make a statement, or they are being pressured by the perpetrator to tell the police officers false information. In these cases, the ability to give a partial score based on inference is ideal (and not possible with the dichotomous ODARA scoring). With the ordinal scoring, the same criteria were used to obtain a full score (i.e., a score of 2), but it allowed for a score of 1 if a reasonable inference of victim fear can be made based on the actions of the victim.

Interrater Reliability

Two graduate students were involved in the data collection for the current research; each coded approximately half of the files that made up the sample. Both student coders received training on how to accurately score and record the ODARA and the coding guide, which included a provision of detailed item descriptions and supervised scoring/coding practice. The ODARA training was delivered by the same person who trained the police service members (who, in turn, was trained directly by the developers of the ODARA). The coding guide training was delivered by the primary researcher (doctoral student). If a coder was familiar with a victim or suspect in a given file, then the file was immediately assigned to the alternative coder in order to minimize bias by ensuring no outside information or case knowledge was available to the coder, to protect the privacy of the complainant and suspect as much as possible, and to minimize conflicts of interest. All coders signed a confidentiality oath in addition to being granted security clearance by the police service prior to accessing police files.

Sixty cases (i.e., 20% of the total sample) were randomly selected to assess the interrater reliability (IRR) of the coding guide and ODARA between the two researchers. The IRR of the coding guide was assessed using Kappa statistics for categorical variables, which provides a robust quantitative measure of the “true” agreement between independent raters. It is calculated by comparing the differences between the level of actual agreement between raters and the level of agreement that would be expected by chance alone (Viera & Garrett, 2005). A commonly cited guideline to Kappa statistic interpretation suggests that Kappa values of .41 to .60 reflect moderate agreement, values of .61 to .80 indicate substantial agreement, and values of .81 to .99 denote almost perfect agreement (1.00 being perfect; Landis & Koch, 1977; Sim & Wright, 2005; Viera & Garrett, 2005). For the current study, a minimum Kappa criterion of $>.61$ was used to ensure substantial agreement between raters. The IRR for all continuous variables (including the dynamic risk items) was assessed using an intraclass correlation coefficient ($ICC_{2,1}$; two-way, random effects model), which is a commonly used statistical approach to assess the consistency of measurements made by independent raters on a continuous variable (Bartko & Carpenter, 1976). Critical values for single measure ICC’s were as follows: $ICC \geq .75 =$ excellent, $.60 \leq ICC < .75 =$ good, $.40 \leq ICC < .60 =$ moderate, $ICC < .40 =$ poor (Fleiss, 1986).

The IRR was tested on the first 20 files after the initial training, with both researchers coding the file without consulting each other or knowing each other’s scores. Fifteen items had inadequate IRR (i.e., 3.6% of the total items in the coding guide), so the police files and training manuals were consulted to identify, discuss, and clarify any discrepancies. A decision was mutually agreed upon for how to score the item, at which

point all previously collected data was corrected. IRR was tested again at the midpoint of data collection and again towards the end of data collection, to ensure that no procedural drift was occurring and that the IRR remained adequate throughout the coding process. At the second IRR check, three variables failed to meet the minimum agreement level. Items were clarified and corrected to produce increased reliability, and one item (the aforementioned escalation item) was eliminated. At the third IRR check, all items met the minimum agreement level. It should be noted the vast majority of categorical items – including all of the ODARA items – had excellent Kappa values (i.e., $k \geq .90$; Sim & Wright, 2005; Viera & Garrett, 2005), and likewise the majority of continuous items has ICC(2,1) values above .90.

Procedure

The goals of the current study were accomplished by examining retrospective police records involving IPV and extracting relevant information as articulated in the coding guide. Permission to access police records for the purpose of the current study was obtained from the Chief of Police for the involved police organization. Prior to accessing these files, researchers involved in the project obtained official security clearance and authorization to access and review relevant files from the police force. Police files were identified and provided by a civilian member of the police force with knowledge on and experience with this process. All information was reviewed on site by the authorized research personnel. A synopsis of the current project was submitted to the Ethics Review Board of the University of New Brunswick, Saint John, and was approved on December 7, 2017 (REB file number 044-2017).

Police records generated for IPV events between June 2011 and February 2016 were coded as index events for the current study. Any subsequent police files for the identified suspects were reviewed for 12 months to determine if any further IPV events were reported to the police (e.g., if the index was in February 2016, then their file was reviewed until February 2017). The ODARA was scored based on file information for the index events and for all subsequent IPV events within the 12-month follow-up period. In addition, the police-scored ODARAs were extracted from the case files whenever available. Recidivism was considered any new incident that meets the same inclusionary criteria used to select the index IPV case, regardless of whether the victim was the same or different from the index event (i.e., because the ODARA captures perpetrator risk and not victim risk).

Results

Case Descriptive Statistics

Most suspects in the current sample were or had been in a heterosexual relationship with the victim (96.1%; $n = 297$). Only 12 were in a same-sex relationship, three of whom were male and nine of whom were female. Over half of the sample were currently in a relationship with the victim at the time of the index event (60.5%, $n = 187$). The other 39.5% ($n = 122$) of the sample were previously in an intimate relationship with the victim. Almost half the suspects (46.6%; $n = 144$) were living with their victims at the time of the index incident. The mean relationship length was 5.1 years ($SD = 6.1$) and ranged from less than one month to 40 years. For those who were no longer in a relationship, the mean length of the separation/dissolution was 11.1 months ($SD = 17.9$) and ranged from less than one month to eight years.

IPV Offender Characteristics

As shown in Table 1, over one-third of the sample had a prior IPV perpetration incident, with male suspects being significantly more likely than female suspects to have such a history, $\chi^2(1, N = 309) = 8.53, p = .003$, Cramer's $V = .27$. That is, forty-two percent ($n = 88$) of male suspects and 25% ($n = 25$) of female suspects had at least one prior charge of IPV on their police record. Likewise, male suspects were more likely than female suspects to have previous charges for at least one pre-index non-IPV violent offense, $\chi^2(1, N = 309) = 12.95, p < .001$, Cramer's $V = .21$, or general (non-violent) offense, $\chi^2(1, N = 309) = 41, p < .001$, Cramer's $V = .36$, on their police record. That is, 73% ($n = 152$) and 79% ($n = 166$) of male suspects had previous non-IPV-related violent and non-violent charges listed in their files, whereas this was true for 52% ($n = 52$) and 43% ($n = 43$) of female suspects, respectively. Female suspects, on the other hand, were significantly more likely than males to have been listed as victims in IPV police files prior to the index event, $\chi^2(1, N = 309) = 37.25, p < .001$, Cramer's $V = .35$. That is, 37% ($n = 37$) of female suspects were previously named as the victim of IPV in a police file, compared to 9% ($n = 18$) of males.

The average researcher-rated ODARA total score was 5.15 ($SD = 2.4$), with 14.6% of the total sample classified as low-risk, 44.7% as moderate-risk, and 40.1% as high-risk. However, there were significant gender differences for both total risk score, $F(1, 305) = 55.814, p < .001, \eta^2 = .15$, and risk classification, $\chi^2(2, N = 309) = 46.06, p < .001$, Cramer's $V = .39$. Male suspects had an average ODARA total risk score of 5.78 ($SD = 2.30$), which was significantly higher than female suspects with an average ODARA total risk score of 3.79 ($SD = 2.03$). In terms of risk level, 9.2% of male

suspects were classified as low-risk, 37.7% as moderate-risk, and 53.1% as high-risk, whereas 26% of female suspects were classified as low-risk, 60% as moderate-risk, and 14% as high-risk. Thus, females tended to have low- to moderate-risk for reoffending, whereas most males were moderate- to high-risk.

Overall, 29.45% ($n = 91$) of the total sample recidivated with at least one IPV-related police contact within the 12-month follow-up period. The rate of recidivism significantly varied by gender, with more males (34%, $n = 71$) than females (20%, $n = 20$) recidivating at least once within 12 months of the index event, but the magnitude of the effect was small, $\chi^2(1, N = 309) = 6.35, p = .01$, Cramer's $V = .14$. There were no gender differences with regards to whether the recidivism involved the same victim, $\chi^2(1, N = 309) = .07, p = .79$, Cramer's $V = 0.3$, with 87% ($n = 79$) of the 91 recidivists targeting the same victim at both the index and the first recidivism. The mean number of post-index, police-reported recidivistic incidents committed by the 91 individuals who reoffended within the one year follow up period was 1.8 ($SD = 1.28$) and ranged from 1 to 7. There were no significant differences in frequency of recidivism between male and female recidivists, $F(1, 89) = 2.59, p = .11, \eta^2 = .28$. Likewise, there were no differences between male and female recidivists in terms of how quickly (in days) they recidivated post-index, $F(1, 89) = 0.29, p = .59, \eta^2 = .003$, with an average of 104 days passed ($SD = 104$) between the index and first IPV reoffence.

Male and female suspects were equally likely to have used intoxicating substances at the time of the index event, $\chi^2(1, N = 309) = 3.03, p = .08$, Cramer's $V = .09$, with 46% ($n = 95$) of males and 35% ($n = 35$) of females having used substances. Weapon use during the index event was equally uncommon for both genders, $\chi^2(1, N =$

309) = .50, $p = .48$, Cramer's $V = 0.04$, with only 14% ($n = 29$) of males and 11% ($n = 11$) of females having used some type of weapon. There also were no gender differences in the type of weapon used, $\chi^2(3, N = 40) = 1.78, p = .62$, Cramer's $V = .21$. Of those who had a weapon, 35% ($n = 14$) used a knife or sharp object, 22.5% ($n = 9$) used a blunt or heavy object, 20% ($n = 8$) used a glass object, and 22.5% ($n = 9$) used something else as a weapon (e.g., pillow, belt, rope, hot curling iron, etc.).

Victim Characteristics

As shown in Table 2, the victims of IPV were predominantly female (70%; $n = 215$). The vast majority were Caucasian (92.6%; $n = 286$), and an additional 2.9% ($n = 9$) were Black. The remaining 4.5% were another ethnicity or it was unspecified in the file ($n = 14$). The mean age of victims was 33 years ($SD = 11.3$). Of those victims who were female, 5.6% ($n = 12$) were pregnant at the time of the index event. Slightly more than half of victims were unwilling to cooperate with the police investigation and expressed that they did not want charges laid against the suspect (53%; $n = 163$). There was a significant difference between male and female victims in terms of non-cooperation with the police investigation, $\chi^2(1, N = 309) = 8.26, p = .04$, Cramer's $V = .36$, with significantly more males (65%, $n = 61$) than females (47%, $n = 102$) choosing not to cooperate. There were no gender differences in victim use of substances at index, $\chi^2(1, N = 309) = .86, p = .35$, Cramer's $V = .05$, with 26.9% ($n = 83$) of victims being under the influence of drugs or alcohol at the time of the index event. Likewise, there were no gender differences in whether the victim sustained a visible injury, $\chi^2(1, N = 309) = .01, p = .93$, Cramer's $V = .01$, with 48.5% ($n = 150$) of the overall sample sustaining a visible injury.

IPV Offense Characteristics

There were no significant differences between male and female suspects in this sample in the types of abuse they committed, as shown in Table 3. Overall, the most commonly perpetrated abuse category (not mutually exclusive) was physical abuse (81.2%; $n = 251$), followed by emotional abuse (55.7%; $n = 172$), property abuse (31.1%; $n = 96$), stalking/harassment (16.5%; $n = 51$), and financial abuse (3.6%; $n = 11$). Table 4 summarizes specific abusive behaviours within each type of abuse as a function of suspect gender (e.g., if the suspect punched, choked, and pulled the hair of the victim, they would have committed three different subtypes of physical abuse). The only abuse type on which males and females significantly differed was physical abuse, $F(1, 305) = 13.11, p < .001, \eta^2 = .04$. Males perpetrated a mean of 2.02 ($SD = 1.58$) different types of physical abuse at the index, whereas females had a mean of 1.37 ($SD = 1.22$).

In terms of the reason for the altercation that led to the index IPV (not mutually exclusive categories), the most commonly cited reason was “issues related to jealousy” (28.8%; $n = 89$), as shown in Table 5. The next most common reason cited was “issues related to children and parenting,” which was somewhat more common for female suspects than for male suspects, $\chi^2(1, N = 309) = 10.99, p = .001$, Cramer’s $V = .19$. Twenty-five percent ($n = 25$) of females and 10.5% ($n = 22$) of males cited children/parenting issues as a reason for their altercation. The next most common reason for the index altercation was “dissolution of the relationship” (11.7%; $n = 36$), which was comparable for males and females. Nearly 7% ($n = 21$) of the cases cited “arguing about their partner’s substance use” as the reasons for the altercation, which also was

comparable for both genders. Although “fighting over money” was an uncommon reason for an IPV incident, it differed for males and females, $\chi^2(1, N = 309) = 4.41, p = .036$, Cramer’s $V = .12$, with 8.1 % ($n = 17$) of males and 2% ($n = 2$) of females citing this reason ($p = .04$ for Fisher’s Exact Test).

Evaluation of Police Use of the ODARA

Description of Subsample

The post-2015 subsample for the evaluation of police use of the ODARA included 155 male suspects and 36 female suspects ($N = 191$). The mean age of these suspects was 34.4 years ($SD = 11.7$). Only 1.7% ($n = 3$) of this subsample were in same-sex intimate partnerships. The post-2015 subsample was predominantly Caucasian (83.8%, $n = 160$), with the remaining 16.3% being Black (6.8%, $n = 13$), Indigenous (4.2%, $n = 8$), another ethnicity (1.6%, $n = 3$), or unknown/unspecified (3.7%, $n = 7$). There were no significant differences in the pre- and post-2015 subsamples in terms of age, $F(1, 307) = 3.15, p = .08, \eta^2 = .01$, or ethnicity, $\chi^2(4, N = 309) = 6.56, p = .161$, Cramer’s $V = .15$. However, the post-2015 subsample differed significantly from the pre-2015 sample in terms of gender, with a smaller ratio of female suspects in the post-2015 subsample, $\chi^2(1, N = 309) = 41.73, p < .001$, Cramer’s $V = .37$. This discrepancy occurred because the post-2015 sample used random sampling, whereas the full sample (expanded to include files prior to 2015) used quota sampling to ensure minimum of 100 female suspects was reached for a different study that was part of the larger research project.

The average police-rated ODARA total score was 5.29 ($n = 118, SD = 2.8$), with 19.3% of the total sample classified as low-risk, 44.7% as moderate risk, and 36.0% as

high risk. However, there were significant gender differences for both total risk score, $F(1, 117) = 16.766, p < .001, \eta^2 = .13$, and risk classification, $\chi^2(2, N = 118) = 19.59, p < .001$, Cramer's $V = .41$. Male suspects had an average ODARA total risk score of 5.66 ($SD = 2.66$), which was significantly higher than female suspects with an average ODARA total risk score of 2.73 ($SD = 1.98$). In terms of risk level, 13.1% of male suspects were classified as low-risk, 47.5% as moderate-risk, and 39.4% as high-risk, whereas 60% of female suspects were classified as low-risk, 26.7% as moderate-risk, and 13.3% as high-risk. Thus, police officers' ODARA ratings tended to rank females as low-risk for reoffending, whereas most males were ranked by officers as being moderate- to high-risk. The mean number of recidivism events in the post-2015 subsample within 12 months did not significantly differ between groups, $F(1, 112) = 3.22, p = 0.08, \eta^2 = .03$. Collapsed across gender, the mean frequency of recidivism events was .22 ($SD = .60$) for low-risk, .59 ($SD = 1.06$) for moderate-risk, and .84 ($SD = 1.24$) for high risk.

Frequency and Accuracy of Police Use of ODARA

Of all the post-2015 index IPV cases ($N = 191$), 62% ($n = 118$) had a police-scored ODARA in the file. However, in only 49% ($n = 94$) of these 118 cases was the ODARA scored in full (i.e., no missing data, total summed accurately, proration completed if applicable). Police-scored ODARAs were significantly less likely to be in the subsequent recidivism files and when they were included, they were even less likely to be scored in full (see Table 6). For calculation purposes, police-scored ODARAs were corrected by converting missing data into zeros and summing the total scores when they were left blank. These corrected scores were used for the subsequent analyses. Of note,

erroneous proration errors were also corrected for, but it did not make a meaningful difference for any of the subsequent analyses so the simpler corrected data (i.e., missing data converted to 0s and total scores summed) were used as they more closely reflect the officer's original ODARA scores.

The intraclass correlation coefficient was calculated to compare interrater reliability of the sum of the ODARA items (i.e., the total ODARA risk score prior to proration) between police and research coders. The ICC for the sum of ODARA items was .61, which is technically is considered “good agreement” (Fleiss, 1986), but is markedly lower than what was expected given previous validation research on the ODARA (typically ranging from .90 to .95; Hilton et al., 2004; Hilton, Rice, & Harris, 2010; Jung & Buro, 2017. Cohen's kappa was also calculated to compare police and researcher ODARA risk level classification, which was poor ($\kappa = .24$). To further understand the source of the less than acceptable agreement, Cohen's Kappa was run to examine the item-level agreement between police raters and researchers. The average per item Kappa was moderate at $\kappa = .44$ (i.e., below the acceptable target of $\kappa = .61$ for substantial agreement), ranging from the lowest agreement of 9% on the ODARA's “victim barrier to support” item (poor agreement) and the highest agreement of 79% on the “assaulted victim while pregnant” item (substantial agreement). Kappa values for the other ODARA items are summarized in Table 7. It should be noted that inconsistencies in scoring often were the result of errors in officers' scoring of information that was available in the case file records (i.e., information equally available/known to the researcher coders and police officers), rather than stemming from information being

known to the officers but not reported in case records for coders to review for their scoring purposes.

A binomial logistic regression was performed to examine several factors that could influence the likelihood that a police officer would complete the ODARA for a given case. Logistic regression analyzes the relationship between multiple independent variables and a categorical dependent variable. Logistic regression expresses the impact of the independent variables in terms of odds (Park, 2013). That is, the odds ratio (OR) is a comparative measure of the probability of an event occurring over the probability of it not occurring. The OR also can be used to determine if exposure to a particular independent variable increases or decreases the likelihood of a particular outcome, and it can compare the magnitude of various independent variables that might influence that likelihood. An OR of 1.00 indicates that an independent variable does not affect the odds of the outcome occurring. An OR greater than 1.00 indicates that exposure to the independent variable is associated with higher odds of the outcome, with each unit increase on the predictor independent variable. An OR less than 1.00 indicates that exposure to an independent variable decreases the odds of the outcome, with each unit decrease on the predictor independent variable (Park, 2013). For example, if the independent variable of gender is coded as 0 (male) and 1 (female), and the OR is -2.5, then the odds of police using the ODARA would be 2.5 times less likely for females than for males. The Wald test was used to determine statistical significance for each of the independent variables within the prediction model.

To determine the sample size required to conduct this logistic regression, the guidelines proposed by Peduzzi et al. (1996) were used. Based on their research, it is

recommended that the formula ' $N = 10 k/p$,' where k is the number of covariates in the equation, and p is the smallest of the proportions of negative or positive cases in the population. In this case, there are seven covariates in the equation, so $k = 7$. It is more difficult to determine p , as there is no known research to inform what proportion of cases can be expected to be positive or negative (i.e., proportion of cases for which the ODARA was used versus not used by police). The current policy in place requires that the ODARA is used for all incidents of IPV for which a chargeable offence has likely occurred, but compliance is unknown. To be conservative, it was hypothesized that police would use the ODARA in approximately half of the cases. This would make $p = 0.5$. The formula then reads: ' $N = (10*7)/ 0.5$.' Using this formula, the sample size required was $N = 140$. Although this analysis was done on only those cases for which the index was dated 2015 or later, there were still 191 eligible cases, thereby exceeding the minimum sample size criteria.

Independent variables examined as part of the logistic regression included suspect gender, victim cooperation with the investigation, suspects' use of a weapon or alcohol, suspect and victim relationship status and living arrangement, and whether there was a visible victim injury. The dependent variable was whether the police used the ODARA at the index event. Analyses were conducted to test the assumptions of the binomial logistic regression, finding that there were no outliers greater than two standard deviations from the mean. In addition, tests of multicollinearity indicated that all terms in the equation had a variance inflation factor (VIF) less than two, which is considered excellent. The logistic regression model was statistically significant, $\chi^2(10, N = 191) = 23.77, p = .008$. The model explained 15.9% (Nagelkerke R^2) of the variance in police

use of the ODARA and correctly classified 68.1% of the cases. Of the seven predictor variables, only two were statistically significant: suspect gender and substance use at index (as shown in Table 8). Police officers were 3.7 times more likely to complete an ODARA for male suspects than a female suspect. They were also 2.1 times more likely to use the ODARA if the suspect was under the influence of alcohol or drugs at the time of the index event relative to when no substance use was present.

Police-Rated Risk and Recidivism

A receiver operating characteristic curve (ROC) analysis was used to assess the predictive validity of the police-rated ODARA, compared by suspect gender. The ROC analysis is a widely used statistic in risk assessment research (Finch et al., 2016; Rice & Harris, 2005; Swets, Dawes, & Monahan, 2000). It was developed based on signal detection theory, whereby a plot is created to demonstrate the true positive detection rate (i.e., percent of offenders correctly classified as recidivists) against the false positive detection rate (i.e., percent of offenders incorrectly classified as recidivists; Hajian-Tilaki, 2013). That is, the ROC analysis plots the tool's sensitivity (hit rate) and specificity (false alarm rate) across various cut points of the scale being tested in the analysis, which produces a curve. As the specificity of the tool increases, the sensitivity decreases and vice versa. A major advantage is that ROC analyses are not influenced by base rates of the target behaviour in a population (Campbell & Messing, 2017). This means that even if a behaviour (in this case, IPV recidivism) occurs infrequently, ROC analyses can provide consistent, accurate evaluations of a risk tool.

The area under the curve (AUC) in a ROC analysis provides an evaluation of how well the risk tool predicts the outcome, ranging from an index of 0 to 1.00. If the

AUC is .50, then the tool's ability to predict outcomes accurately is equal to chance, whereas an AUC of 1.00 represents perfect prediction (Park et al., 2004). In the current study, the AUC value represents the probability that a randomly selected recidivist would score higher than a randomly selected non-recidivist. Rice and Harris (2005) suggested that AUCs of .556 to .638 represent a small effect size, whereas .639 to .713 suggest a moderate effect, and .714 or higher reflects a large effect size. These are commonly used rules of thumb and were used in the proposed study to interpret the AUCs.

DeLong et al. (1988) have developed a commonly accepted method of comparing AUCs for situations in which ROC analyses are used as a metric of outcome prediction. This method involves a simple comparison of 95% confidence intervals of each AUC value, whereby non-overlapping or just touching confidence intervals represent different population parameters (i.e., significantly different). Confidence intervals that overlap by approximately one quarter or less also represent two different population parameters and are considered significantly different. Confidence intervals that overlap by more than one quarter represent similar predictive validity of the tool for both groups and are, therefore, not significantly different (Campbell et al., 2007).

There were four cases for which there was a known follow-up period of zero days that were filtered out for these analyses ($N = 114$; i.e., fled the province, immediate detention without bail for at least 12 months, death). The ROC analysis indicated that the police-rated ODARA total score was not predictive of IPV recidivism beyond chance, $AUC = .58$, 95% CI [.47, .69] (see Figure 1). When separated by gender, a large effect size was found for the prediction of recidivism among females, $AUC = .71$, 95%

CI [.47, .96], whereas the predictive accuracy for recidivism among males corresponded to a small (near chance) effect size, AUC of .53, 95% CI [.42, .65]. Of note, the CI width for female suspects was very large and overlapped with .50 (chance prediction). In addition, the CI's overlapped by more than 25% for males and females, suggesting they were sampling the same population parameter. Therefore, the AUC for females is still considered to show poor predictive accuracy using the police-rated ODARA scores (see Table 9, for a comparison of police and researcher AUCs and CIs).

A binomial logistic regression was also performed to examine the ability of police risk scores to predict recidivism after controlling for gender. The binomial logistic regression assumptions were tested first. The results indicated that the police-rated ODARA score had a linear relationship with the logit of the dependent recidivism variable, as assessed by the Box-Tidwell (1962) procedure and using a Bonferroni correction resulting in statistical significance being accepted when $p < .02$ for the test of linearity (Tabachnick & Fidell, 2014). There was one outlier with a standardized residual value of 3.72 standard deviations. Further inspection revealed that this outlier was an individual who had the highest rate of reoffending in the entire sample (i.e., they were the suspect in seven additional IPV-related police contacts within 12 months). Although they were an outlier in their rate of reoffending, the case is still relevant as part of the risk distribution and, therefore, the case was retained for further analyses. Finally, tests of multicollinearity indicated that all terms in the equation had a VIF less than two, which is excellent. However, the logistic regression model was not statistically significant, $\chi^2(3, N = 114) = 5.68, p = .13$. None of the predictor variables were statistically significant. Thus, consistent with the AUC finding, the ODARA as

completed by police was not predictive of recidivism within the 12-month follow-up period, nor was suspect gender.

The relationship between police scored ODARAs and recidivism was further explored through a Kaplan-Meier survival analysis (Kaplan & Meier, 1958). More specifically, a survival analysis statistic was used to examine whether suspects' police-scored ODARA risk classification could predict the time (in days) to recidivism. Survival analyses are used to model the time from an initial event to the next occurrence of the event, with the amount of time between events known as the "survival time" (Luke & Homan, 1998; Parmar & Machin, 1995). In a sense, time to an event is the outcome variable in survival analysis (Yoshihama & Horrocks, 2003). The Kaplan-Meier survival analysis is a nonparametric procedure that can analyze several predictors (including binary outcome measures, such as recidivism versus non-recidivism). A coefficient that quantifies the effect of the covariates on the hazard of IPV (i.e., recidivism) can be estimated, and exponentiation of the coefficients produces a "hazard ratio," which is an estimate of relative risk. A hazard ratio of 1 would indicate that the predictor variable has no effect on time to IPV recidivism.

As for police-scored ODARA risk classifications, low-risk offenders had an average time to first recidivism of 326.17 days ($SD = 21.27$) or approximately 47 weeks, collapsed across gender (i.e., there were no significant gender differences). This was longer than the moderate- and high-risk offenders, who had an average of 272.32 days ($SD = 19.11$; approximately 39 weeks) and 279.84 days ($SD = 23.05$; approximately 40 weeks), respectively. A log-rank test was conducted with a survival analysis to determine if there were differences in the survival distribution as a function of ODARA

risk classification levels identified by police-scored ODARAs. The survival distributions for the three risk categories were not statistically significantly different from each other, $\chi^2(2, N = 114) = 3.42, p = .18$, nor was there a difference for male and female suspects in this sample. Thus, the risk level categories produced from police-scored ODARAs were unable to discriminate between suspects on time passed before the occurrence of a new police contact as a suspect of IPV.

Overall, the ROC analysis, logistic regression, and survival analysis all similarly indicated that police-scored ODARAs were not accurately predicting occurrence of, nor the time to, recidivism within 12 months in this sample. The same three analyses were conducted using the researcher-rated ODARA scores for comparison purposes, using the same subsample. Using the researchers' scores, 8.8% ($n = 10$) of the suspects were classified as low-risk (down from 19.3% using police scores), 41.2% ($n = 47$) were classified as moderate-risk (comparable to the 44.7% using police scores), and 50% ($n = 57$) were classified as high-risk (up from 36% using police scores). The ROC analysis indicated that the researcher-rated ODARA total score was predictive of IPV recidivism with a moderate effect size collapsed across gender, $AUC = .69, 95\% CI [.60, .76]$ (see Figure 2). When separated by gender, a large effect size was found for the prediction of recidivism among females, $AUC = .79, 95\% CI [.64, 1.00]$, whereas the predictive accuracy for recidivism among males corresponded to a moderate effect size, AUC of $.64, 95\% CI [.60, .74]$. However, as with the police-rated ODARA scores, the CI's overlapped by more than 25% for males and females, suggesting they were sampling the same population parameter (i.e., the predictive validity of the researcher-scored ODARA was not significantly different for male and female suspects).

The logistic regression model, which examined the ability of researcher-rated ODARA scores to predict recidivism after controlling for suspect gender, was statistically significant, $\chi^2(3, N = 114) = 12.63, p = .005$. The model explained approximately 15% (Nagelkerke R^2) of the variance in recidivism and correctly classified 71.9% of cases. Gender was not a significant predictor of recidivism, but the researcher-rated ODARA total score was a unique predictor of recidivism above and beyond gender. Every one-point increase on the ODARA represented 1.33 times higher odds of recidivism.

In terms of days to recidivism by researcher-scored risk classification, low-risk offenders had an average time to first recidivism of 341.3 days ($SD = 22.48$). The moderate-risk offenders had an average of 325.51 days ($SD = 14.12$), and high-risk offenders had the shortest time to recidivism with an average of 242.97 days ($SD = 20.61$). The Kaplan-Meier survival analysis indicated that there was a significant difference in time to recidivism as a function of researcher-rated risk classification groups, $\chi^2(2, N = 114) = 10.59, p = .005$. Pairwise log-rank comparisons were conducted to determine which risk groups had different survival distributions. A Bonferroni correction was made with statistical significance accepted at the $p < .0167$ level. There was a statistically significant difference in survival distributions between the high-risk group and both the moderate-risk group, $\chi^2(2, N = 114) = 8.13, p = .004$, and the low-risk group, $\chi^2(2, N = 114) = 3.35, p = .05$. However, the survival distributions for the low- and moderate-risk groups were not statistically significant from each other, $\chi^2(2, N = 114) = 0.29, p = .59$. Once again, there was no meaningful difference for male and female suspects for any of these survival analyses or post-hoc comparisons.

Police-Rated Risk and Risk Management

The average number of actions taken by police officers after responding to an IPV incident (post-2015 files; $N = 191$) was 9.04 ($SD = 3.69$). However, the average number of police actions taken was significantly higher when the IPV call involved male suspects ($M = 9.52$, $SD = 3.70$) as opposed to female suspects ($M = 6.97$, $SD = 2.86$), $F(1, 189) = 14.98$, $p < .001$, $\eta^2 = .07$. As reported in Table 10, the pattern of increased police action for cases involving male suspects (and typically involving female victims) remained similar when separately considering official criminal justice actions (i.e., protocol), $F(1, 189) = 4.91$, $p = .028$, $\eta^2 = .03$, and victim-directed discretionary police actions, $F(1, 189) = 20.08$, $p < .001$, $\eta^2 = .10$. However, cases involving male suspects received significantly fewer suspect-directed discretionary police action than female suspects, $F(1, 189) = 11.22$, $p = .001$, $\eta^2 = .06$.

A hierarchical multiple regression was conducted to determine whether the addition of police-rated ODARA total risk scores improved the prediction of the number of risk management strategies used by officers (i.e., police action), over and above gender alone. To assess the assumption of linearity, a scatterplot graph was plotted of the total ODARA risk score against average number of police action variables with a superimposed regression line. A visual inspection of this plot indicated that a linear relationship between these variables exists. There was independence of residuals, as assessed by a Durbin-Watson statistic of 2.03. A visual inspection of the scatterplot also indicated that there were no major outliers. Casewise diagnostics were examined for outliers $> \pm 2$ standard deviations from the mean, finding none. There was homoscedasticity, as assessed by visual inspection of a plot of standardized residuals

versus standardized predicted values. Residuals were normally distributed as assessed by visual inspection of a normal probability plot. Thus, it was appropriate to proceed with the regression analysis.

Using the cases of the post-2015 subsample for which a police-scored ODARA was available ($n = 114$), suspect gender was entered in Block 1 of the hierarchical multiple regression analysis (exceeding sample requirement of $n = 88$, as indicated by the a priori power analysis for a medium effect (.15), a power of .90, and an alpha of .05). Gender significantly contributed to the prediction of the number of police actions taken, $R^2 = .06$, $F(1, 112) = 6.68$, $p = .01$, explaining 5.6% of the variance in number of risk management strategies used. The police-rated ODARA score was entered at Block 2 and significantly explained an additional 5.9% of the variance in the model, $\Delta R^2 = .06$, $F(2, 111) = 7.22$, $p = .001$.

As a follow-up to the above analyses, chi square analyses were conducted to examine specific police actions (e.g., arrested suspect, recommended pretrial detention, etc.) by police-rated risk classification level (i.e., low, moderate, and high). Only four police actions varied significantly across ODARA risk categories, such that as risk level increased, police were more likely to conduct these actions (i.e., “risk-response match”): looked up suspect’s criminal record, $\chi^2(2, N = 114) = 8.70$, $p = .02$, Cramer’s $V = .43$, arrested suspect, $\chi^2(2, N = 114) = 9.76$, $p = .01$, Cramer’s $V = .38$, recommended charges be laid, $\chi^2(2, N = 114) = 8.02$, $p = .03$, Cramer’s $V = .35$, and recommended suspect be held in pretrial detention, $\chi^2(2, N = 114) = 14.45$, $p < .001$, Cramer’s $V = .38$. Although not statistically significant, most of the other police actions showed a trend towards increasing in likelihood as risk level increased (see Table 11).

There were no gender differences in risk-response match for the majority of the police actions; however, the risk-response match was significant for male suspects for both arrest, $\chi^2(2, N = 99) = 7.31, p = .02$, Cramer's $V = .37$, and pretrial detention, $\chi^2(2, N = 99) = 8.02, p = .01$, Cramer's $V = .38$, such that as male suspects' risk increased, so did the likelihood that police would make an arrest or recommend pretrial detention. On the other hand, risk-response match was not significant for female suspects when it came to arrest, $\chi^2(2, N = 15) = 2.5, p = .29$, Cramer's $V = .41$, and pretrial detention, $\chi^2(2, N = 15) = 2.14, p = .34$, Cramer's $V = .38$. That is, police officers were equally unlikely to make an arrest or recommend pretrial detention to the Crown for low-, medium-, and high-risk female suspects.

Risk, Risk Management, and Recidivism

The relationship between risk score, risk management, and recidivism were examined two ways: first with a mediation model and then with a moderation model. The first method used will be the Sobel-Goodman tests for mediation (Preacher & Hayes, 2004, 2008), using the Hayes (2018) adaption. A mediation analysis is a statistical technique that allows for the testing of how a causal variable, X , has an effect on an outcome variable, Y , through its effect on a third variable, M , and M 's effect on Y . In this case, the causal variable is suspect risk level (X ; i.e., as scored by police officers on the ODARA), the outcome variable is IPV recidivism (Y ; i.e., as recorded in police files within 12 months of the index event), and risk management is the mediation variable (M ; i.e., the number of police actions taken). Ideally, all variables in a mediation model would be continuous variables, which would allow regression analyses to be used to test each relationship in this model. It is also common for the independent variable X

to be a categorical variable, as long as the mediator and outcome variable are both continuous.

Various scholars have developed ways to examine mediation effects when the outcome variable is dichotomous, which typically involves using a moderation analysis as a follow-up analysis to increase one's confidence in the results (Hayes, 2018; Iacobucci, 2012; MacKinnon & Dwyer, 1993; MacKinnon et al., 2007; Williams & Stansfield, 2017). The mediation approach can be calculated with a formula developed by MacKinnon and Dwyer (1993), using software developed by Hayes (2018). This process involves utilizing a logistic regression between X (continuous risk score) and Y (binary recidivism variable), ordinary least squares regression analysis between X and M (continuous risk management variable), and another logistic regression between M and Y . The equations are then rescaled using what are called "variance adjusters," which allows for comparable standard errors and parameter estimates (explained in detail by Iacobucci, 2012). This is the approach that was used in the current study. A priori power analysis indicated that a sample of 99 was required to detect a significant mediation with a medium effect (.15), a power of .90, and an alpha of .05.

Prior to conducting the mediation analysis, the assumptions of the regressions were tested. These assumptions were all met, except for one outlier that was greater than two standard deviations from the mean risk score. The case was the same individual previously discussed (who had a high rate of reoffending), so they were retained in the sample for this analysis. In the ordinary least squares regression model, risk score was significantly related to greater number of police actions taken, $b = .36$, $SE = .11$, $p < .001$, 95 % CI [.15, .57]. The logistic regression model included risk score and risk

management as predictors of recidivism. The results demonstrated that police-rated risk score was not significantly independently associated with recidivism, $b = .13$, $SE = .08$, $p = .12$, $OR = .11$, 95% CI [-.03, .29], nor was police action significantly independently associated with recidivism, $b = -.08$, $SE = .07$, $p = .27$, $OR = .05$, 95% CI [-.22, .06]. The indirect effect was tested using non-parametric bootstrap confidence intervals derived from 5000 samples. The results demonstrated that the indirect effect (IE) coefficient was also not significant, $IE = -.03$, $SE = .03$, 95% CI [-.10, .02]. This finding indicates that the relationship between police-rated ODARA risk scores and recidivism was not mediated by the number of risk management strategies used (there were no meaningful gender differences).

To ensure robustness of any findings from the mediation analysis, a logistic regression analysis was used to investigate whether police-rated risk score will moderate the effects of risk management on recidivism, as recommended by Williams and Stansfield (2007). Using Peduzzi et al.'s (1996) formula for determining sample sizes for logistic regression, $N = (10*k)/p$, with $k = 2$ and $p = .2$, a sample size of 100 is required (p is based on Moser's (2012) finding that approximately 20% of a IPV perpetrators in a New Brunswick sample had subsequent IPV-related police contact within 12 months). Both predictors (risk score and risk management) were centered around their means (Aiken & West, 1991) before computing the interaction term, and all terms were entered into the model together. Assumption tests indicated that the police-rated ODARA score and the risk management variable had a linear relationship with the logit of the dependent recidivism variable, as assessed by the Box-Tidwell (1962) procedure and using a Bonferroni correction resulting in statistical significance being

accepted when $p < .02$ (Tabachnick & Fidell, 2014). There were no outliers greater than two standard deviations from the mean. Tests of multicollinearity indicated that all terms in the equation had a variance inflation factor (VIF) less than two, which is considered excellent. The results indicated that the logistic regression model was not statistically significant, $\chi^2(3, N = 114) = 3.39, p = .34$ (similar for both male and female suspects). None of the predictor variables, including the interaction term used to test the moderation effect, were statistically significant. This result is consistent with the findings of the previous mediation analysis.

Development and Evaluation of the Dynamic Supplement to the ODARA

The second and third goals of the current study were to develop a dynamic supplement to the ODARA and then to conduct a preliminary evaluation of the effectiveness of this new tool in comparison to the ODARA. The full sample ($N = 309$) was used for all the following analyses, except for 11 cases that were filtered out as they were not available for the 12-month follow-up period due to known death, departure from the province, or incarceration immediately following the index event. This final sample ($N = 298$) was made up of 66.8% male suspects ($n = 199$) and 33.2% female suspects ($n = 99$). All of the analyses involved in developing the dynamic risk supplement were examined separately for male and female suspects, and there were no notable gender differences in most cases. The results are only reported separately for males and females when there was a meaningful difference to report.

The results of the previously reported interrater reliability checks indicated that there was consistently poor IRR for one of the dynamic risk factors, namely ‘recent escalation in IPV’ item. Multiple efforts were made to clarify the operationalization of

this variable to improve the reliability of the scoring of this item; however, police officers rarely recorded in their notes if there was an indication of escalation in IPV and the researchers did not have access to case notes from previous incidents in a suspect's file. Therefore, the reliability of this item did not improve enough to incorporate it in analyses. As such, the escalation item was eliminated from further examination.

Kendall's Tau-b correlations were run to determine the relationship between individual dynamic risk factors ('relationship problems,' 'excessive jealousy or control,' 'employment or financial problems,' 'suicidal/self-harming behaviours,' 'antisocial attitudes,' 'substance use during/leading up to current IPV,' and 'victim fear of future IPV') and the dichotomous IPV outcome (i.e., recidivism versus non-recidivism) to test Douglas and Skeem's (2005) first criterion for dynamic risk factors. There are no guidelines for determining the strength of the association for Kendall's Tau-b, the values can range from -1 to +1, and values nearing zero equal weak or no association between the two variables. Of the seven dynamic risk factors, only 'antisocial attitudes'⁴, $\tau_b = .13, p = .02$, was weakly positively associated with IPV recidivism within the 12-month follow-up period (for male and female suspects alike). None of the other dynamic risk factors were significant, with Tau-b values ranging from -.04 to .10 (see Table 12).

⁴ This analysis was re-run using the "split file" feature of SPSS in order to examine the effects of suspect race on the relationship between the 'antisocial attitudes' variable and the recidivism variable. The sample was comprised of 85.6% Caucasian suspects ($n = 255$), 3% Indigenous suspects ($n = 9$), 5% Black suspects ($n = 15$), 2.7% suspects of an "other" racial/ethnic background ($n = 8$), and 3.7% ($n = 11$) suspects of unknown racial/ethnic background (i.e., the information was not tracked in the police file). This is approximately representative of the population demographics of the community from which the data were drawn, according to the 2016 census data (note: the source is unreferenced to maintain anonymity of the sample police service). For the purpose of the follow-up analyses, the 11 individuals with no race data available were filtered out and the remaining individuals were categorized as "BIPOC" ($n = 32$) or "Caucasian" ($n = 255$). The results of the split Kendall's Tau-b analysis demonstrated that the 'antisocial attitudes' variable continued to have a significant positive correlation with recidivism for Caucasian suspects, $\tau_b = .15, p = .01$, but it was not significant for BIPOC suspects, $\tau_b = .01, p = .97$.

To examine whether any of the dynamic risk factors were associated with specific types of abuse outcomes (rather than simply examining recidivism versus non-recidivism), composite variables were computed for emotional abuse, physical abuse, property abuse, stalking/harassment, and financial abuse. If at least one of the behavioural indicators of each of these abuse types was coded positively at the first recidivism event, then the composite variable was coded as “yes” with a value of one (1). If no behavioural indicators were present, then the composite variable was coded as “no” with a value of zero (0). Kendall's Tau-b correlations were computed, finding mostly weak or no relationships between the dynamic risk factors and the five IPV abuse types of recidivism (Table 12). There was only one dynamic risk factor that had a moderate strength relationship with two different IPV outcomes, which was ‘victim fear of future abuse.’ Increased victim fear was moderately positively associated with the ‘stalking/harassment’ IPV recidivism outcome, $\tau_b = .26, p < .01$, and the ‘physical abuse’ outcome, $\tau_b = .24, p < .01$ (for both male and female suspects).

Survival analysis was used to examine whether suspects’ ratings on the dynamic risk variables could predict the time (in days) to recidivism. Each dynamic risk factor was rated as not present (0), partially present (1), or present (2) for each suspect in the sample. Log-rank tests were conducted to determine whether there were significant differences in the survival distribution for the different levels of each dynamic risk factor. As summarized in Table 13, these findings indicated that the ‘antisocial attitudes’ variable was the only one out of the seven dynamic risk variables that showed a significant difference in the survival distributions for the three levels of the variable,

$\chi^2(2, N = 298) = 5.83, p = .05$.⁵ Suspects who were rated as having “no evidence” of antisocial attitudes had an average of 308.90 days to recidivism ($SD = 10.20$). Those rated as having “partial evidence” of antisocial attitudes had an average of 291.91 ($SD = 18.97$) days to recidivism. Finally, suspects rated as having “evidence” of antisocial attitudes had an average of 266.65 ($SD = 12.84$) days to recidivism. Pairwise log rank comparisons were conducted to determine which levels of the variable had different survival distributions. A Bonferroni correction was made with statistical significance accepted at the $p < .017$ level. There was a statistically significant difference in survival distributions for time to recidivism for the “no evidence” and “evidence” levels of antisocial attitudes, $\chi^2(2, N = 298) = 5.78, p = .016$ (for both male and female suspects). However, the survival distribution for the “partial evidence” level was not statistically significantly different from the “no evidence” level, $\chi^2(2, N = 298) = .97, p = .33$, or the “evidence” level of antisocial attitudes, $\chi^2(2, N = 298) = .61, p = .44$.

To evaluate criterion two of the conceptualization of a dynamic risk factor, the Reliable Change Index (RCI) analysis, developed by Jacobson and Truax (1991), was used to determine whether the dynamic risk factors have the ability to fluctuate over time. The RCI can assess whether change on a given item is reliable (i.e., not explained by chance or measurement error). The idea behind this analysis is that all measurement is unreliable to some extent and that a difference between two scores could be due to measurement error rather than, for example, a true and reliable change in dynamic risk (Morley & Dowzer, 2014). It is calculated based on the individual’s score at Time 1,

⁵ A follow-up survival analysis split by suspect race demonstrated that the ‘antisocial attitudes’ item continued to predict the time in days to recidivism for Caucasian suspects at all three levels of the variable, $\chi^2(2, N = 255) = 6.91, p = .03$, but it was not significant for BIPOC suspects, $\chi^2(2, N = 32) = .16, p = .92$.

their score at Time 2, and the standard error of the difference between these two scores. If the person recidivated more than once in the follow-up period, then the first recidivism event was used for their Time 2 score. The change is considered reliable, or unlikely to be the product of measurement error, if the RCI was greater than ± 1.96 (Ogles, Lunnen, & Bonesteel, 2001). Because RCI analyses are based on individual change, it can be conducted using extremely small sample sizes (Zahra & Hedge, 2010). Time 2 data was only available for those who recidivated; therefore, there was a reduced sample size for the RCI analyses ($n = 87$).

The first step to calculate the RCI involved computing the difference between each dynamic risk factor at the index and first recidivism event for each suspect in the sample and calculating the standard deviation for each person's difference score. The standard deviation value was then used to calculate the standard error for each dynamic risk factor, for each person in the sample. After this, the RCI was calculated by dividing the difference value by the standard error for each person. Summarized in Table 14, the results of the RCI analyses indicated that each of the seven dynamic risk factors demonstrated the ability to reliably fluctuate over time.⁶ There were no differences in reliable change based on suspect gender for any of the dynamic risk factors examined.

Internal Validity

The reliability of the dynamic risk items as a total scale was examined by calculating the Cronbach's alphas and intercorrelations. First, a composite variable for the total dynamic risk scale was calculated by summing the ratings of each suspect on the seven dynamic risk factors (ranging from 0 to 2 at the item level, and 0 to 14 at the

⁶ There was no difference for Caucasian and BIPOC suspects in RCI for the 'antisocial attitudes' variable.

total scale level; $M = 5.11$, $SD = 2.88$). Reliability is typically considered the extent to which a group of items measure the same construct, as evidenced by how well they vary together. High internal consistency allows the researcher to interpret the composite score as a reliable measure of the underlying construct (Henson, 2001). Assuming the dynamic risk variables effectively measure underlying risk for IPV, moderate to high estimates of internal consistency would be expected.

Cronbach's alpha can measure the internal consistency of several items by measuring the homogeneity of the items (Litwin, 2003). It is essentially a measure of how well items complement each other in their measurement of different aspects of the same underlying construct. Cronbach's alphas range from 0 to 1.00, with values nearing 1.00 representing higher internal consistency. A common guideline for interpreting this measurement is to be cautious about items with an alpha below .70, and to consider deleting items that would increase the alpha if dropped (McMillan & Schumacher, 2001). The internal consistency analysis revealed a Cronbach's α of .51 for the total scale of the dynamic risk factors, indicating poor internal consistency. Item-total statistics indicated that none of the dynamic risk factors were disproportionately contributing to poor internal consistency (i.e., there were no items that, if deleted, would improve the Cronbach's Alpha). Clark and Watson (1995) suggest that if items measure the same latent construct, then the item-total correlations should be greater than .50, although others recommend a more lenient benchmark of .30 (Cristobal et al., 2007). Using the Clark and Watson (1995) criterion, none of the dynamic risk items had acceptable internal consistency with the overall scale. Using the more lenient criterion, the only two items that had acceptable item-total correlations were 'victim fear' and

‘antisocial attitudes,’⁷ both with $r = .38$ (although the strength of association was somewhat higher for male suspects, $r = .39$ than female suspects, $r = .29$), for the ‘victim fear’ item). The other item-total correlations ranged from $r = .12$ (for ‘alcohol or drug problems’) to $r = .29$ (for ‘employment/financial problems’), which were below the cut-off for acceptable item-total correlations and did not vary by suspect gender.

The internal consistency also can be examined using intercorrelations between items within the scale (i.e., inter-item; DeVellis, 2003). Assuming items measure the same latent construct, then items within the scale should be positively correlated. Clark and Watson (1995) suggest that the average inter-item correlation should fall between .15 and .50. Tests of multicollinearity identified weak inter-item correlations for five of the seven dynamic risk variables, with average inter-item correlations of .06 for ‘alcohol/drug use,’ .10 for ‘suicide/self-harm,’ .11 for both ‘jealousy/controlling’ and ‘relationship problems,’ and .14 for and ‘employment/financial problems.’ The ‘victim fear’ and ‘antisocial attitudes’ items were the only two dynamic risk factors that met the criterion for acceptable multicollinearity, both with average inter-item correlations of .19. The average inter-item correlation for all seven items was .13, indicating weak overall multicollinearity (no differences were observed based on suspect gender).

Predictive and Incremental Validity

Given the lack of internal validity, the dynamic risk composite variable was not used in subsequent analyses (however, an interested reader can see the results of the originally planned analyses with the composite variable in Appendix C). There were

⁷ There was no significant change in the results when the internal consistency was examined split by race, such that the ‘victim fear’ and ‘antisocial attitudes’ variables remained the only two dynamic risk factors that met the minimum criteria for acceptable multicollinearity for both Caucasian suspects ($r = .39$ and $r = .31$, respectively) and BIPOC suspects ($r = .40$ and $r = 3.7$, respectively).

only two items that were statistically significantly correlated with recidivism outcomes within the 12-month follow-up period, which were ‘antisocial attitudes’ (correlated with general IPV recidivism) and ‘victim fear’ (correlated with IPV recidivism that involved stalking/harassment or physical abuse). Given that ‘victim fear’ is already a dynamic item included on the ODARA, ‘antisocial attitudes’ was the only item used in subsequent analyses to examine predictive and incremental validity.

A receiver operating characteristic curve (ROC) analysis was used to assess the predictive accuracy of ‘antisocial attitudes’ for imminent (< 1 month), short-term (up to 3 months), medium-term (up to 6 months), and long-term (up to 12 months) time periods. That is, the ‘antisocial attitudes’ item score for the index offence was used to predict any recidivism during these time points. Using the full sample, the ROC analysis indicated that ‘antisocial attitudes’ was not predictive above chance for IPV recidivism within one month, AUC = .59, 95% CI [.49, .69], three months, AUC = .60, 95% CI [.51, .68], six months, AUC = .59, 95% CI [.52, .67], or twelve months, AUC = .58, 95% CI [.51, .65].⁸ The DeLong et al. (1988) method of comparing AUCs was used to compare the produced AUCs for male and female suspects. The results of these analyses, as displayed in Table 15, demonstrated no significant gender differences in the predictive validity of the ‘antisocial attitudes’ item.

⁸ The follow up ROC analyses for the ability of the ‘antisocial attitudes’ item to predict IPV recidivism at multiple time points showed the general trend of a higher AUC value for Caucasian suspects and a lower AUC for BIPOC suspects, but the effect sizes remained small across all time points. That is, the ‘antisocial attitudes’ variable had small effect sizes in prediction of IPV recidivism for Caucasian suspects within one month, AUC = .63, 95% CI [.53, .73], three months, AUC = .63, 95% CI [.54, .72], six months, AUC = .61, 95% CI [.52, .68], or twelve months, AUC = .59, 95% CI [.52, .67]. The ‘antisocial attitudes’ variable was not predictive above chance for IPV recidivism for BIPC suspects within one month, AUC = .46, 95% CI [.13, .77], three months, AUC = .44, 95% CI [.17, .72], six months, AUC = .57, 95% CI [.35, .78], or twelve months, AUC = .50, 95% CI [.30, .71].

The predictive validity of the ‘antisocial attitudes’ item collapsed across gender had a weak effect size across all time points, whereas the researcher-scored ODARA had a moderate effect size across all four time points. Using the DeLong et al. (1988) method, the average overlap of the 95% confidence intervals for the ‘antisocial attitudes’ item and the ODARA total score was 25.23%, indicating that the ODARA was stronger at predicting IPV recidivism across time points (see Table 16). Next, the incremental validity of adding ‘antisocial attitudes’ to the ODARA was examined. The dynamic item was recoded to be consistent with the scoring of the ODARA (i.e., dichotomous), such that ratings of zero (no evidence of antisocial attitudes) were recoded as zero (not present) and ratings of one (partial or inconclusive evidence) and two (evidence of antisocial attitudes) were recoded as one. A Kendall’s Tau-b analysis was performed again to examine the relationship between the dichotomously scored ‘antisocial attitudes’ item, and the results were very similar to the original variable, $\tau_b = .14, p = .02$.⁹ Table 17 displays the results of the DeLong method comparison of the ODARA and the ODARA plus the dichotomously scored ‘antisocial attitudes’ item (i.e., “ODARA + AA”). The ODARA + AA had moderate predictive validity at each time point, and although the AUC improved for each of the four recidivism periods, the overlap of the 95% Confidence Intervals exceeded 25%, demonstrating that the addition of the ‘antisocial attitudes’ item did not add significant predictive value above and beyond the original ODARA total score.

⁹ Likewise, the split file results of the Kendall’s Tau-b analysis demonstrated that the dichotomously scored ‘antisocial attitudes’ variable continued to have a significant positive correlation with recidivism for Caucasian suspects, $\tau_b = .16, p = .01$, but it was not significant for BIPOC suspects, $\tau_b = .02, p = .92$.

A Cox regression survival analysis was used to examine whether the dichotomous ‘antisocial attitude’ could provide incremental validity for predicting time to recidivism (in days) beyond the ODARA total score. A cox regression is a semi-parametric procedure that can analyze several predictors (including binary outcome measures, such as recidivism versus non-recidivism). Unlike parametric regression models, Cox regression does not assume a particular distribution for the outcome variable, but it does assume that the effect of a time-fixed covariate on the outcome variable is proportionate (Yoshihama & Horrocks, 2003). A coefficient that quantifies the effect of the covariates on the hazard of IPV can be estimated, and exponentiation of the coefficients produces a “hazard ratio,” which is an estimate of relative risk. For the current study, the covariates were the researcher-scored ODARA total score (entered in Block 1) and the ‘antisocial attitudes’ item score (entered in Block 2). A priori power analysis indicated that a sample of 88 was required to detect a significant result with a medium effect (.15), with a power of .90 and an alpha of .05, which was surpassed with an $N = 298$ for this analysis. The ODARA risk classification, in Block 1, significantly contributed to the prediction of time to IPV recidivism, $\chi^2(1) = 18.89, p < .001$. In Block 2, although the overall model still significantly predicted time to recidivism, $\chi^2(2) = 19.11, p < .001$, the addition of the ‘antisocial attitudes’ item did not demonstrate incremental predictive improvements beyond the ODARA, ($\beta = .22, p = .35, \text{Exp}(B1) = 1.25$), $\chi^2(1, N = 298) = .88, p = .35$.¹⁰

¹⁰ Follow-up Cox Regression analyses demonstrated that the model was similarly predictive of time to IPV recidivism at Block 1 for BIPOC suspects, $\chi^2(1) = 5.31, p = .02$. However, in Block 2 in the model no longer significantly predicted recidivism for BIPOC suspects, $\chi^2(2) = 5.10, p = .08$ after the addition of the ‘antisocial attitudes’ item, ($\beta = -.19, p = .76, \text{Exp}(B1) = .85$), $\chi^2(1) = .09, p = .76$. Of note, the a priori power analysis indicated that a sample of 40 was required; therefore, this analysis was underpowered with $n = 32$.

The final comparison between the ‘antisocial attitudes’ item and the ODARA involved examining the added value of this dynamic item in detecting change in risk over time. The previously discussed RCI analysis indicated that any change on the ‘antisocial attitudes’ item was reliable (i.e., not explained by chance or measurement error). To examine incremental validity, the same steps were used to calculate the RCIs for the both the ODARA and the ODARA + AA, which were compared to determine if the addition of the ‘antisocial attitudes’ item increased the tool’s sensitivity to reliable change over the 12-month follow-up period to change. As with the prior RCI analysis, only those who recidivated were able to be included in the analysis; therefore, the sample was reduced to $N = 87$. If the person recidivated more than once in the follow-up period, then the first recidivism was used for their Time 2 score. The change was considered reliable, or unlikely to be the product of measurement error, if the RCI was greater than ± 1.96 (Ogles, Lunnen, & Bonesteel, 2001).

The results of the RCI analyses for the researcher-scored ODARA indicated that only 14.3% ($n = 13$) had a reliable decrease in their risk score, 38.5 % ($n = 35$) had no reliable change, and 47.3% ($n = 43$) had a reliable increase in their risk score. As a follow-up to the RCI analyses, a paired-samples t-test was used to determine whether there was a significant mean difference between the total ODARA scores for suspects at index and at the first recidivism event. The results indicated that the suspects in this subsample (i.e., the recidivists) had a lower average ODARA score at the index ($M = 6.01$, $SD = 2.23$) then at the first recidivism event ($M = 6.74$, $SD = 1.65$), 95% CI [-1.03, -.42], $t(90) = -4.69$, $p < .001$. A chi-square test of independence was conducted between ODARA risk classification and direction of change using the RCI analyses results (i.e.,

reliable decrease, no change, reliable increase). There was a statistically significant association between ODARA risk classification at index and the direction of change in score on the ODARA between the index event and the first recidivism, $\chi^2(4, N = 126) = 34.33, p < .001$, Cramer's $V = .43$. Post hoc testing was conducted using adjusted standardized residuals and a Bonferroni correction was applied, resulting in statistical significance being accepted when $p < .0056$ (Tabachnick & Fidell, 2014). As displayed in Table 18, all three risk classification groups (i.e., low-, moderate-, and high-risk) were most likely to have a reliable increase in ODARA risk score over time, and the moderate- and high-risk groups were second most likely to have no change in risk score.

The RCI analyses for the ODARA + AA demonstrated that 30.8% ($n = 28$) of participants in the subsample of recidivists had a reliable decrease in their ODARA + AA score, whereas 25.3% ($n = 23$) had no change and 44% ($n = 40$) had a reliable increase in their ODARA + AA score between the index event and the first recidivism event. The follow-up paired samples t-test indicated that the suspects in this sample had a lower average ODARA+AA score at the index ($M = 6.68, SD = 2.46$) than at the first recidivism event ($M = 7.25, SD = 1.85$), 95% CI [-0.94, -1.19], $t(90) = -3.02, p = .003$. A chi-square test of independence was conducted to examine change in ODARA + AA as a function of ODARA risk classification at the time of the index. There was a significant association between risk classification and the direction of change in the ODARA + AA score between the index event and first recidivism, $\chi^2(4, N = 126) = 21.06, p < .001$, Cramer's $V = .34$. Bonferroni corrected post hoc testing, displayed in Table 19, demonstrated that the higher the risk classification score, the greater the likelihood of a reliable decrease in ODARA + AA score.

As other researchers have done (Frerichs & Tuokko, 2006), RCI scores for the ODARA and ODARA + AA were converted into a dichotomous variable (presence or absence of reliable change), which, in turn, was analyzed using chi-square tests of significance to compare them. The Chi-square analysis revealed that there was a significant difference between the ODARA and the ODARA + AA in terms of the presence or absence of reliable change, $\chi^2(1, N = 126) = 20.60, p < .001$, Cramer's $V = .49$. This indicates that the addition of the 'antisocial attitudes' item to the ODARA increases its sensitivity to change in risk over time.¹¹

Discussion

Police across Canada are becoming increasingly involved in risk assessment and risk management when responding to cases of IPV (Public Safety Canada, 2015). As such, the current research set out to examine the implementation of the ODARA, a well-established actuarial IPV risk-assessment instrument that has been adopted by police agencies in New Brunswick. One of the primary objectives of the current dissertation was to examine how officers use the ODARA in a frontline context, including how frequently they use the tool, factors that influence their likelihood to complete the tool, the reliability of their scoring, and the effectiveness of the tool to predict risk and guide police decision-making. Problems with officer scoring limited the utility of the ODARA, the implications of which are discussed below. Prior to reviewing the results of the current dissertation, a discussion of gender is provided to offer context to these findings. Another main objective of the current research was to develop and evaluate a dynamic risk supplement to the ODARA to increase the tool's sensitivity to change in risk over

¹¹ As with the previous RCI analyses involving the 'antisocial attitudes' item, there was no difference in reliable change between Caucasian and BIPOC suspects.

time, ability to identify the criminogenic needs of an IPV suspect, and overall utility to guide police decision-making. Only the ‘antisocial attitudes’ dynamic risk item significantly added to the ODARA, so subsequent analyses focused on its incremental contributions. Implications for policing and future research are discussed below.

The Role of Gender

The role of gender is a very important consideration in the study of IPV. In terms of the IPV that gets reported to the police in Canada, approximately 70% of cases involve female victims and male perpetrators (Buczycycka, 2018). In the portion of the current sample that was drawn as a random sample (i.e., the post-2015 data), approximately 80% of cases involved male suspects and female victims. This percentage was consistent with the ratios found in a previous study conducted within the same jurisdiction as the current study (Moser, 2012), as well as other Canadian samples (Peterson & Schroeder, 2019). However, we know from self-report data that a large portion of male victims of IPV do not call the police, and this skews the perception of IPV as a predominantly male-perpetrated issue (Statistics Canada, 2016). It is important for research on IPV and risk assessment to examine female perpetrators (with male victims) in their samples; however, this specific research foci was the subject of another project using the current dataset. It is also important to examine IPV between same-sex partners, but unfortunately this was not possible in the current dissertation because there were only three same-sex couples in the sample. Although a detailed examination of the role of gender in IPV was outside the scope of the current dissertation, gender was examined as part of the analyses of the current study whenever possible. The most

relevant results are discussed next to offer context to the current data as it relates to the evaluation of the ODARA in a frontline policing context.

In terms of suspect characteristics, there were far more similarities between males and females than differences. Notably, there were no differences between male and female suspects in the types of abuse they committed, with physical abuse being the most common type reported to police in the current sample. These results resemble Moser's (2012) findings, who also found that physical abuse was the most common type of IPV perpetrated by both male and female offenders. When specifically examining physical forms of violence in the current sample, males perpetrated more diverse types of physical abuse than females. However, there were no gender differences in the level of injury sustained by the victim, with nearly half of incidents resulting in a visible, physical injury. This finding is also consistent with other Canadian samples (McCormick et al., 2011; Moser, 2012). In terms of suspects' histories of IPV, an examination of police records found that male suspects were much more likely than female suspects to have been charged with IPV prior to the index event (42% and 25%, respectively). This finding again supports the notion that, when it comes to police-reported IPV, there are meaningful gender differences.

Police Response to Male and Female Suspects. The fact that there tended to be such strong gender differences in the pre-index criminal records of IPV suspects in the current sample may help explain why police officers seemed to respond differently to the men and women with whom they interacted. For instance, despite gender-inclusive provincially sanctioned definitions of IPV and ODARA training that included instruction to use the tool with both male and female IPV suspects (New Brunswick

Multicultural Council, 2018; A. Totten, personal communication, December 13, 2017; Province of New Brunswick, 2012), officers in the current (post-2015) sample were nearly four times more likely to complete the ODARA when suspects were male.

Officers in this (post-2015) sample took significantly more official criminal justice response actions against the suspect in IPV cases when the suspect was male. This finding could be at least partially mitigated by the fact that, on average, police officers rated male suspects as higher risk on the ODARA than female suspects, and they tended to use more risk management strategies for higher risk suspects. However, there are several reasons why this explanation is incomplete. First, officers completed an ODARA in only 62% of the cases in the current sample, meaning in 38% of the cases their decision-making could not have been informed by a risk estimate. Second, when the ODARA was completed, officers' decision to arrest or recommend pretrial detention corresponded with risk classification for male suspects but not for female suspects. In fact, they were less likely to arrest a high-risk female suspect than a low-risk male suspect. Third, there were significant differences in police and researcher scores on the ODARA by suspect gender. For instance, although police officers rated 60% of females as low-risk, researchers' retrospective coding of these same files indicated that only 11% of female suspects in this (post-2015) sample were low-risk. Therefore, it is likely that gender expectations and biases contributed to both officers' scoring of the ODARA and their risk management choices.

Aspects of victims' experience with police in IPV situations may differ in some ways for males and females as well, and differentially influence police response. For instance, police in the current sample were significantly more likely to offer services and

other discretionary aid to victims who were female. Often victim-directed discretionary aid involved giving victims information relevant to their situation, such as: a) how to apply for civil procedures (e.g., small claims court, family court for divorce/custody issues), b) how to obtain legal aid, c) how the criminal justice system works (e.g., police can recommend charges even without the victim making a statement), d) where they can receive support and free services in the community, and e) information about IPV (e.g., the cycle of violence, the results of the ODARA). Victim-directed discretionary actions also included making referrals to the Victim Services Unit or community-based IPV services, developing a safety plan with the victim, bringing them to a safe house/shelter or other safe place to stay the night, following up with the victim in the subsequent days, confirming and/or contacting a victim support person, establishing a police contact person for the victim, making a referral or personally helping improving security factors at the victim's residence (e.g., calling the landlord to have locks changed, referring a security company, helping fix a broken door), and other actions (e.g., calling pet services to have their animals taken care of).

Not only were female victims in the current (post-2015) sample more likely to receive discretionary aid compared to male victims, but female suspects also were more likely than the male suspects to receive discretionary aid from police officers. Suspect-directed discretionary aid included following up with the suspect post-incident to check in, referring them to community services (e.g., mental health services, addiction services, anger management, agencies that help with affordable housing and childcare), or other actions. These differences in police action as influenced by victim and suspect gender is comparable to the research by Storey and Strand (2013), through which they

found the average number of police-recommended risk management strategies increased with higher police-rated risk scores on the B-SAFER for male suspects but not for female suspects.

Comparably, Storey and Strand (2017) prospectively examined police officers' risk management recommendations for IPV victims and perpetrators after completing the B-SAFER and found that officers were more likely to base their risk management recommendations on assessed victim vulnerability factors and the level of risk posed by the perpetrator when the victim was female and the perpetrator was male (Storey & Strand, 2017). Russell and Sturgeon (2019) conducted a study that similarly examined police decision-making in response to heterosexual and same-sex partners in an IPV incident. In their study, 309 American police officers read a hypothetical IPV scenario with variations in the suspect and victim genders. They found that officers believed that non-arrest options, such as providing informal advice and mediation, were fairer for suspects who were gay males compared to heterosexual males. Similarly, they thought it was fairer to ask one of the involved parties to leave the premises to cool off when the perpetrator was a gay male or heterosexual female. Russell and Sturgeon (2019) concluded that police officers may consider IPV perpetrated by gay men and heterosexual women as less serious.

Research findings such as these, in which differences in police decision-making related to suspect and victim gender occur, may tap into police perception of male victims as being less deserving or in need of help, or perhaps a perception of female perpetrators engaging in acts of IPV as a reaction to their own victimization. Unfortunately, this latter perception might be based in some degree of truth when it

comes to police-reported IPV, as evidenced by the fact that over one-third of the female suspects in the current sample were previously listed as victims in police records of IPV cases (i.e., 39%), compared to only about one-tenth of the male suspects (i.e., 9%). Therefore, it might be expected to see an increased effect of gender-bias in police response in situations where there was mutual IPV occurring between both partners (i.e., to see the female-perpetrated IPV as ‘self-defence’). However, these cases did not meet the inclusionary criterion that the suspect and victim must be unambiguously identified by the police officer in the file; therefore, these cases were not included in this sample. Another possibility is that female victims receive more police support due to disproportionate rates of cooperation with the police. For instance, in the current (post-2015) sample, female victims were significantly more likely to cooperate with police investigations (e.g., provide a statement). In contrast, male victims were less likely to cooperate with the police and more likely to state that they did not want charges laid against the suspect. A possible interpretation is that police may be less likely to offer discretionary aid to victims who do not cooperate with their investigations, perhaps assuming that if a victim does not want to avail of the official criminal justice system resources, then they would not want to avail of other resources. Further research is required to examine this hypothesis.

Gender Differences in Risk and Recidivism. The male suspects in the current sample had higher average risk scores than the female suspects (according to researcher-scored ODARAs), and, therefore, unsurprisingly had higher rates of recidivism. However, there were no differences between male and female recidivists in the frequency of their recidivism, quickness to recidivism, or the level of injury their victim

incurred. The current study also found similar predictive validity effect sizes to those reported by Moser (2012) and Hilton et al. (2014) using researcher-scored ODARAs, with moderate to large effect sizes for both male and female suspects. In addition, the examination of dynamic risk factors in the current study did not find meaningful gender differences in the predictive validity in any of the dynamic items. These results are consistent with the findings of Dutton et al. (2006), who concluded from a literature review that female and male IPV abusers share many of the same characteristics and risk factors. Similarly, Medeiros and Straus' (2006) literature review found that there were no differences in IPV risk factors for male and female perpetrators in 72% of cases. On the other hand, the literature used to identify and study risk factors for IPV, including for the current dissertation, is almost entirely based on samples of male IPV offenders and, therefore, could be missing important risk factors that are specific to females (Ménard et al., 2009).

Summary of the Role of Gender. Overall, the current research suggests that female and male perpetrators of IPV are just as likely to reoffend and they may share many of the same risk factors. It also suggests that the ODARA may be able to predict IPV recidivism, frequency of recidivism, and quickness to recidivism equally for females and males; however, the literature on use of the ODARA with female suspects is scarce and more research with diverse and larger samples is required to further evaluate such use. In addition, the ODARA only has value if it is completed consistently and correctly. The current research found that police appear to have gender biases in their completion and implementation of the ODARA. That is, the current results suggest that police officers are more likely to complete the ODARA with male suspects, to

incorrectly classify female suspects as low-risk and male suspects as high-risk, to over-intervene with male suspects, and to offer more discretionary aid to both female victims and female suspects. Findings such as these suggest that the cultural effects of hegemonic masculinity and heteronormative understandings of IPV may continue to penetrate police services in Canada, which emphasizes a need to continue to highlight such issues in the training of police officers.

Police Use of the ODARA

Frequency and Predictors of ODARA Use

Despite increased use of IPV risk assessment by police services across Canada, there is very little research on how these tools are implemented by frontline officers (Bourgon et al., 2018; Department of Justice, 2018b; Grant & Rowe, 2011; Kropp, 2008; Humphrey et al., 2005; Northcott, 2012). The current study offered some initial insights into frontline compliance with the implementation of the ODARA, which is the IPV risk assessment tool that has been endorsed by the New Brunswick Association of Chiefs of Police and the Province of New Brunswick (Office of the Chief Coroner, 2015). Results demonstrated that although the majority of the randomly selected IPV files in the current sample included a police-scored ODARA, approximately 38% did not. Furthermore, if an ODARA was included in the index file, it was significantly less likely that a new ODARA would be completed in subsequent files involving the same suspect.

Out of a variety of possible predictor variables examined, only suspect substance use at the index event and suspect gender significantly predicted whether the officer used the ODARA. The influence of suspect substance use on police officers' decisions to use the ODARA is likely due to its salience as a risk factor. At the broadest level, in

Canada there is a widespread societal belief that alcohol use causes aggression (Government of Canada, 2006; Field et al., 2004). Individual acceptance of this belief may increase a police officer's likelihood to attend to this variable on the scene and to respond accordingly. In addition, officers are likely to be aware of the fact that substance use is well-established by research as a risk factor for general and violent crime (i.e., in addition to IPV; Pernanen et al., 2002; Room et al., 2005). They are also likely to have previous experiences with intoxicated suspects that could influence how they respond in such cases. For instance, previous research has documented that substance-impaired suspects are more likely than their sober counterparts to resist arrest, to assault a police officer, and to use a weapon against a police officer (Alpert et al., 2006; Garner & Maxwell, 1999). These beliefs, knowledge, and experiences could prime police officers to pay attention to substance use when they arrive on the scene for an IPV event, and perhaps heighten their perception of risk in those situations.

The fact that officers in the current study were more likely to use the ODARA for male suspects is not particularly surprising given that their training historically instructed them to use the tool in cases of male-on-female IPV, which was the original intent of the ODARA, and because males were more commonly represented as suspects. It is also possible that the decision to use the risk tool may be tied to officers' decisions of arrest, as police officers in the current study were significantly more likely to arrest male suspects. In addition, an arrested IPV suspect's police file is required to include a completed ODARA by protocol for submission to the Crown Prosecutor for charge approval and bail decision making in New Brunswick (Office of Attorney General, 2017; Office of the Chief Coroner, 2015). However, compliance with this policy is not

known, and the timing of completion of the ODARA was not coded; as such, it is unclear whether the ODARA was used to inform the police officers' decisions or completed after the fact to fulfil a requirement for file management.

Overall, of the post-2015 case records, a little over one-third of the IPV files in the current sample did not include a police-scored ODARA at all, despite the fact that these cases clearly met criteria as IPV. In addition, approximately one-third of cases had only partially completed police-scored ODARAs on record. Thus, of eligible files, only one-third contained a fully completed police-scored ODARA. This rate is far from full compliance with the policing standards endorsed in New Brunswick, which state that a validated risk assessment is to be used to guide decision-making in response to IPV calls (Public Safety Canada, 2017). It remains unclear as to why officer compliance with policing standards was not higher, but some experts in the field speculate that frontline officers are often able to negotiate their way out of even the most specific policies based on their discretionary power on a case-by-case basis (Barton, 2003; Paolone, 2003). Evidence for this use of discretion can be found in the plethora of studies which have demonstrated that even when clear policies are implemented to mandate the arrest and investigation of all IPV suspects, full police compliance with these policies has never been reached (Belknap & Hartman, 2000; Buzawa, 1982; Cole & Gertz, 1998; Ho, 2000; Simpson et al., 2006).

A limitation of the current study is that it did not examine police officers' characteristics, experiences, or attitudes about IPV risk assessment protocols in order to explore how these factors may affect their use of the ODARA. These examinations were beyond the scope of the current dissertation and the data available. However, theory and

research are available in the existing literature on IPV that can provide context for these findings. For instance, the theory of planned behaviour is a commonly referenced and examined theory in research on police decision-making (Ishoy, 2016). This theory specifies that (a) an individual's attitude about a behaviour predicts behavioural intent, and thereby predicts behaviour; (b) social pressures to engage in a certain behaviour increase the likelihood that an individual will do the behaviour in question; and (c) an individual's perception of their ability, obligation, and behavioural control in a situation will prevent or increase their likelihood to do the behaviour (Ajzen, 1991). The theory of planned behaviour could certainly be used to provide a framework to examine police use of the ODARA. For instance, this theory might posit that ODARA use could be explained by the officer's attitudes about the ODARA and intent to use it, the perceived social norms from their coworkers and supervisors regarding ODARA use, and the degree to which the officer feels they are obligated and able to complete and implement the tool. There is a variety of research on police decision-making that could help contextualize the current finding that one-third of the IPV files in the current sample did not include a police-scored ODARA. These research findings can be understood within the framework of the theory of planned behaviour, reflecting the themes of (a) personal attitude, (b) social pressures, and (c) perception of control.

Personal Attitude. In terms of individual attitude or "buy-in," a variety of previous qualitative research has examined police attitudes towards IPV risk assessment tools and evidence-based policing in general. In a review and narrative analysis of 73 studies from Canada, the United States, and the United Kingdom on risk assessment and risk management amongst various types of professionals, Viljoen et al. (2018) concluded

that if there is no buy-in for use of risk assessment tools, then they do not get adequately used by frontline police officers. One study from New Brunswick found that officers appreciate that risk assessment helps structure data gathering in their investigations, but fear that risk assessment tools could encroach on their discretionary power (Ballucci et al., 2017). This has been found in a variety of other samples as well (Barton, 2003; Campbell et al., 2017; Humphrey et al., 2005). Another common finding in police research is that officers tend to have a negative attitude about and resistance to any tasks that they believe will add to their already heavy paperwork loads (Flanagan, 2008). During the present study, officers also anecdotally reported frustration with the added paperwork that completing the ODARA creates.

Social Pressures. Available research also supports the idea from the theory of planned behaviour that social pressures within a police service to engage in a certain behaviour will increase its likelihood. For example, prior research has found that supervisors play an important role in shaping the attitudes and behavioural norms for officers under their command, and subsequently their compliance with assessment protocol (Grant & Rowe, 2011; Humphrey et al., 2005). On the other hand, some research indicates that social pressures from external agencies are much less influential. For instance, Kalyal (2019) found that Canadian frontline officers are more likely to resist tasks associated with evidenced-based policing when they originate from third party researchers or organizations, believing that civilians are not able to understand police work. Similar findings have been found in various other samples as well (Lum et al., 2012; Palmer, 2011; Palmer et al., 2019; Telep & Lum, 2014; Wood et al., 2014). Similarly, in a qualitative study using interview data with 38 police executives from

across Canada, Koziarski and Kalyal (2020) found that frontline police officers are more likely to resist elements of evidence-based policing (such as risk assessment tools) when they feel far-removed from the research process and are not consulted during development and implementation of new projects.

Overall, these findings indicate that police services should work to build supervisor buy-in to the ODARA to help create an atmosphere of positive social pressure to comply with ODARA policies. In addition, it suggests there may be a need for researchers to increase their effort to consult and collaborate with police services at all stages of research and development in order to increase agency level buy-in to evidence-based policing strategies and tools (including use of the ODARA), as this may be more effective at influencing behaviour than social pressures from civilian third parties. For example, this collaboration could involve having police officers serve on a research committee to help with identifying important research questions and goals, developing methodologies, assisting with data collection, offering input on interpretation of results, co-creating recommendations and future directions, and collaboratively disseminating the results.

Perception of Control. In keeping with the theory of planned behaviour's notion that the perception of control influences behaviour, Ballucci et al. (2017) also found that many New Brunswick police officers were concerned that completing an IPV risk tool would be a waste of time if the Crown prosecutor does not consider the suspect's risk score in their own decision-making. Anecdotally, this view is consistent with concerns raised by police officers around the metaphorical watercooler throughout the process of data collection for the current study. The fear or belief that the ODARA would not be

considered by the Crown may cause officers to feel they have no control over the implementation of the ODARA, and thereby could reduce the likelihood that an officer will complete an ODARA. This could be potentially addressed by clarifying to officers in training that the ODARA has value beyond Crown consideration. Ideally, ODARA training should include instruction on risk prediction (scoring), risk management (focusing on application by police officers directly), and an overview of the RNR model/evidence-based policing. The intention is that increased education and training will help officers understand the direct benefits of IPV risk assessment for policing, facilitate justified resource allocation, and help alleviate police frustration.

In addition to increased officer education, Crown prosecutors ideally should be taking police-scored ODARAs into consideration during their decision-making processes for charge approval, bail, prosecution, and sentencing. This is not an unrealistic possibility. For instance, the Department of Public Safety issued a statement in 2015 stating that the Public Prosecution Services in New Brunswick was overseeing the training of all Crown Prosecutors on the ODARA, and that prosecutors have been advised that the completed tool will form part of all investigation files relating to IPV (Office of the Chief Coroner, 2015). In addition, the Public Prosecution Services Operational Manual for IPV stipulates that prosecutors must take any risk assessment forms in the file into consideration, particularly when determining whether to release a suspect and, if so, which conditions of release to issue (Office of Attorney General, 2017). It is unclear, therefore, if it is simply police officers' perceptions or if it is a reality that Crown Prosecutors are not taking a suspect's ODARA score into consideration in practice. Increased communication between these sectors about risk

information may elucidate the use of the ODARA in guiding decision-making by the Crown, although increased education and training of the Crown and judges on how to use the ODARA to inform legal arguments and decisions may be required as well.

There is a variety of other research that supports the idea from the theory of planned behaviour that a police officer will be less likely to engage in a behaviour (e.g., completing the ODARA) if they do not feel in control of what to do and how to do it. For instance, Koziarski and Kalyal (2020) found that frontline Canadian police officers are more likely to resist evidence-based policing tasks if they seem abstract or have not been operationalized into actionable tasks. Similarly, Grant and Rowe (2011) found evidence to suggest that officers who do not fully understand the importance and benefit of IPV risk assessment are far less likely to follow related protocols. Conceivably, these challenges could be a barrier with the application of the ODARA in the current sample if officers did not have an adequate understanding of how to practically use the ODARA beyond scoring it and possibly presenting to the Crown. In fact, other than to suggest sharing the results of the ODARA with the IPV victim, the ODARA manual does not offer prescriptive steps or practical suggestions for how to apply the results of the tool to guide decision-making (Mental Health Centre Penetanguishene Research Department & Ontario Provincial Police Behavioural Sciences Section, 2005).

To improve frontline use of the ODARA, police services should consider increasing the emphasis in the ODARA training on the practical ways that police officers can use risk assessment to structure their investigation, detect risk, guide their decision-making, mediate recidivism, and ultimately increase victim safety. These efforts should involve a digestible review of relevant research findings that are presented

in a manner that emphasizes the relevance and realistic nature of evidence-based policing. Future research should examine IPV risk assessment policy compliance of police officers before and after training and education of this nature. In addition, a possible next step to increase buy-in and use of the ODARA could involve employing a focus group of police officers to brainstorm and develop a guide for possible risk management strategies (i.e., actionable tasks) that correspond with various levels of ODARA risk.

Reliability of Police Scoring of the ODARA

Although it is important to understand and improve police compliance with IPV risk assessment protocol, it is possibly even more important to demonstrate the quality of its implementation (Helmus, 2018). Previous research has established that police officers can reliably score the ODARA in controlled research contexts (Hilton et al., 2004); however, it is important that research examines the ability of officers to accurately score the ODARA in a frontline context. A few studies have examined frontline use of the ODARA in other countries (Switzerland, Australia, and New Zealand), and they have generally found support for the ODARA in these contexts, but some have shown notably lower effect sizes than what is obtained in controlled research contexts (Gerth et al., 2017; Lauria et al., 2017; Nimmo, 2012). The variations in effect sizes could be due to a variety of factors, such as the fact that their samples were from different countries and cultural contexts (Australia, New Zealand, and Switzerland). Unfortunately, these scarce preliminary studies have focused on the application of the tool and/or its predictive validity rather than examining officer scoring. Because the accuracy of police officers' scoring was not examined in previous research, issues with

scoring cannot be ruled out as a possible factor affecting the utility of the ODARA in these field studies.

The results of the current study found several problems with the accuracy of the ODARA scoring by officers. First, approximately only half of the police-rated ODARAs were scored in full (i.e., no missing data, total summed accurately, and proration completed if applicable). In addition, although the interrater reliability between the two graduate student researchers for the ODARA items was excellent, the interrater reliability between the researchers and police officers was poor. Only four out of 13 items met the minimum criteria for acceptable interrater reliability, including ‘confinement of the partner at the index,’ ‘victim or offender have more than one child,’ ‘victim has a child with a different partner,’ and ‘offender assaulted the victim while pregnant.’ Interestingly, all five of the ODARA items that are scored using the suspect’s police and criminal record files had particularly poor interrater reliability between officers and researchers (e.g., whether the offender has a history of domestic assault or other violence). Often police officers marked these items as “unknown” or left them blank, indicating that they may not be consulting the suspect’s record when completing the ODARA. Thus, valuable risk information may be missing from their consideration of risk appraisals and associated risk mitigation response. Although previous research has not directly evaluated frontline police scoring accuracy, some studies have found higher predictive validity of officer risk scores when officers must submit their risk ratings to their supervisors for review prior to finalizing them (Belfrage & Strand, 2012; Storey et al., 2014). Given this, police services could consider using supervisor consultation, review, or oversight as a possible manner to increase adherence to the

scoring protocol of the ODARA. Future research should examine the possible effects of increased supervisory oversight to decrease avoidable errors such as these.

The item with the weakest interrater reliability between researchers and front-line police officers was ‘victim barriers to support.’ Officers frequently scored this item to indicate that there were no victim barriers to support, despite the fact that there was information in their notes that provided evidence to the contrary. Previous (unpublished) research on police use of the ODARA in New Brunswick has similarly found poor reliability for this item (M.A. Campbell, personal communication, May 19, 2021). This inaccuracy in scoring may be at least partially due to the fact that the electronic scoring form used by officers in the frontline context contains a single statement reflecting this item, which may cause them to forget the more nuanced operational definition of the item. Specifically, officers were prompted on screen with only the statement “Victim has at least one barrier to support, 0 or 1.” However, in the ODARA training and manual, this item is operationalized as “Victim has at least one barrier to support if they: have at least 1 child who is less than 19 years old that they care for; have no phone; have no access to transportation; are geographically isolated; consumed alcohol/drugs before or during the index IPV; and/or have a history of substance use problems, including frequent use of alcohol to the point of intoxication, use of any street/illegal drug, use of marijuana without a prescription, or abuse of medication” (Mental Health Centre Penetanguishene Research Department & Ontario Provincial Police Behavioural Sciences Section, 2005). Although officers would have been taught the full description of each ODARA item in their training, it could be difficult to remember all these nuances when completing the ODARA in a frontline context. The simple “victim

barriers to support” prompt provided to police officers at the time of completing the ODARA may be misleading and cause officers to under-rate this item. In fact, Hilton et al. (2007) previously found that police officer’s scoring of the ODARA was significantly more accurate when they were provided with the scoring instruction booklet rather than just the simple scoring sheet. Their work was an early validation study conducted in a controlled research setting, but it suggests that access to the full instructions/item descriptions could improve officer’s scoring of the ODARA in frontline contexts as well. Police services should consider providing easy access to physical scoring instruction booklets or electronic versions of the same (e.g., clickable drop down items that reveals the full item description).

In addition to accessing the full instructions when completing the ODARA, police officers may benefit from participating in regular refresher training on the ODARA (Viljoen et al., 2018). In the current sample, all officers received the ODARA training at some point prior to the index data collection. However, the length of time between their training and the completion of the ODARA used in the current research was not recorded or examined. Given the extensive duties of a police officer, it is reasonable to expect that they may forget some of the specific details involved in scoring the ODARA over time. Hilton et al. (2007) examined the effectiveness of a one-day training on ODARA scoring with a variety of professionals working in victim services, law enforcement, offender programming, and family services. They found that those who participated in a training and were provided with a scoring instruction booklet (compared to a control group without the training and instruction booklet) had significantly improved scoring accuracy on the ODARA. A limitation of Hilton et al.’s

(2007) study is that they examined post-training scoring on the same day that training was delivered, so the information would have been fresh in the participants' minds. Future research should examine procedural drift by examining officers' ODARA scoring accuracy at various time points following their initial training to help determine how frequently re-training is required.

RNR Evaluation of Frontline Application of the ODARA

As police services across Canada become increasingly involved in risk management and crime prevention, the RNR model has been recognized by Public Safety Canada as a crucial way to prevent crime (Public Safety Canada, 2015, 2017; RCMP, 2019). As such, the frontline use of risk instruments should be evaluated through the lens of the RNR model to determine the degree to which their use informs risk management actions and recidivism outcomes. One of the primary tenants of the RNR model is that a reliable means of appraising risk for recidivism should be used, and that the intensity of the risk management intervention applied should be consistent with the individual's risk for recidivism (i.e., the risk principle). This principle was used in the current study to guide expectations for risk mitigation that could stem from use of the ODARA. To evaluate the degree to which use of the ODARA was consistent with the risk principle, the predictive validity of the ODARA as used by frontline police officers was examined, followed by an analysis of how officers used the risk estimation to guide the risk management strategies used. Finally, the potential mediation effects of police risk management on the relationship between assessed risk level and recidivism was examined. It was hypothesized that this mediation would be significant, based on the theoretical underpinnings of the RNR model and the past research that supports it.

Various meta-analyses have concluded that rates of reoffending are reduced when risk estimates are used to guide risk management (e.g., Hanson et al., 2009; Koehler et al., 2013; Lipsey, 2009; Lowenkamp et al., 2006; Prendergast et al., 2013).

Risk Prediction. The limited preliminary research on the ODARA as used by frontline police officers is generally promising in the field, with the effect sizes being moderate in range (AUCs ranging from .61 to .68; Gerth et al., 2017; Lauria et al., 2017; Nimmo, 2012). However, none of these studies have used a Canadian sample, nor have they specifically examined the reliability of the officers' scoring as done in the current study. Given the unreliable (i.e., undependable) scoring by officers found in the current study, it was not surprising that police-scored ODARAs were not able to predict IPV recidivism above chance for male or female suspects. In contrast, when the ODARA was scored by the researchers using the same file information available to these officers for the same IPV cases, the tool was moderately predictive of subsequent police contact for male- and female- perpetrated IPV and had a comparable effect size to what has been found in previous research (Messing & Thaller, 2013; Moser, 2012; Olver & Jung, 2017; VanderPut et al., 2019). This finding suggests that the problem was not with the tool or the sample, but rather the identified scoring issues in the police-rated ODARAs limited the tool's utility (i.e., the tool is only as good as the quality of the information imputed into it).

Risk Management. Even though the ODARA as scored by police officers in the current sample was not predictive of IPV recidivism within the 12-month follow-up period, it is still useful to examine whether these officers used their knowledge of the suspect's ODARA risk score to inform the risk management strategies they employed.

Afterall, there is little point to risk estimation if that information is not utilized to guide decision-making (Viljoen et al., 2018). In the ODARA training that police officers received, they were taught to match the suspect's level of estimated risk to the intensity of their intervention, so it was hypothesized that this alignment would be found in the current sample. This hypothesis was supported given that as police-rated risk scores increased, so did the number of risk management strategies implemented by the responding officer. Specifically, officers were significantly more likely to look up the suspect's record, issue an arrest, recommend charges, and recommend pretrial detention for moderate- and high-risk suspects compared to low-risk suspects. These findings suggest that officers may have been influenced by the suspect's ODARA risk score to guide their decision-making. However, because the timing of the ODARA completion (relative to other risk management decisions) was not recorded, it is also possible that officers filled out the ODARA after their other interventions as a way to justify their decision-making. Either way, these findings provide preliminary evidence that it would be possible for police to use the ODARA to improve risk-response match. Of course, these efforts are in vain if risk estimation is inaccurate. Future research should seek to replicate these analyses in other samples.

Although this is the first time police application of the ODARA has been examined in a Canadian sample, a few studies have found this same result in police samples from other countries with the ODARA and using other IPV risk tools. For example, Belfrage and Strand (2008) found that the intensity of response to IPV suspects by Swedish police officers increased with the presence of more risk factors as rated on the B-SAFER. One study examined review board decisions at three Canadian jails,

finding that reviewers with access to an offender's risk score from an objective risk classification instrument were significantly more likely to identify low-risk inmates for placement in a halfway house (53.5% identified), compared to reviewers who were blind to the offender's risk score (16% identified; Bonta & Motiuk, 1990). Likewise, an American study compared the fit of young offenders' risk for recidivism and the intensity of services they received, finding that the probation officers were significantly more likely to make service referrals based on risk-need fit after they were trained on and implemented a validated risk assessment tool (Vincent et al., 2012).

Risk Mitigation. The RNR model was originally developed and used for risk reduction in the post-conviction correctional service/supervision context, so the vast majority of the literature to date has focused on that application (Nicholls et al., 2013). However, some research has begun to examine the RNR model as applied to IPV risk assessment in policing contexts, including two Swedish studies that found risk management mediated recidivism when informed by empirically sound risk estimates. That is, they found that the use of a higher volume of risk management recommendations corresponded with *decreased* recidivism in high-risk offenders and over-use of risk management recommendations *increased* recidivism in low-risk offenders (Belfrage et al., 2012; Storey et al., 2014). Although it was hypothesized that the findings of Belfrage et al. (2012) and Storey et al. (2014) would be replicated using this Canadian sample, the current study found that risk management intensity did not mediate the risk-recidivism relationship, despite the fact that police officers used a suspect's ODARA risk score to inform their risk management response.

The failure to find significant mediation for risk mitigation should not be viewed as a failing of the ODARA's impact because this is likely a function of the unreliable scoring of the ODARA by officers and the subsequent poor predictive validity of IPV recidivism in the current sample. The reliability of police scoring was not examined in Belfrage et al.'s (2012) prospective study of the SARA, nor in Storey et al.'s (2014) prospective study of the B-SAFER, but in both studies these authors did note that officers had to review their SARA/B-SAFER ratings with a supervisor prior to filing their risk management recommendations. This level of accountability may have increased their attention to detail in completing the respective IPV risk assessment tools and thereby may have ensured more reliable scoring. Both studies did examine the predictive validity of the police-rated risk scores, and unlike the current study, found that officer's risk ratings were predictive of recidivism. The obvious conclusion here is that unreliable scoring makes for a risk tool of little value, both in terms of its predictive validity and the effectiveness of decisions stemming from that erroneous risk estimate.

Development and Evaluation of a Dynamic Supplement to the ODARA

Using a validated risk assessment tool to estimate risk is an important step in applying the RNR model, although integrating risk management strategies that address identified dynamic risk factors is a step that is often neglected in policing contexts (Viljoen et al., 2018). The ODARA is well equipped as a second-generation risk tool to provide a reliable estimate of risk, but an unfortunate disadvantage of the ODARA is that it tends to emphasize static risk. That is, many of its items are based on historical factors that, once indicated, cannot be undone or reduced; therefore, the ODARA has been criticized for being less useful for informing imminent risk, measuring fluctuations

in risk, or identifying risk management strategies that target dynamic risk factors to promote change (Hoge & Andrews, 1996; Kebbel et al., 2019).

One of the original goals of the current research was to develop a supplement to the ODARA that has a greater emphasis on dynamic risk for IPV recidivism while staying true to the intent of the ODARA as an efficient tool that can be easily scored with information that is readily accessible to police officers. Several possible dynamic risk factors were examined for inclusion in the dynamic supplemental tool, including ‘escalation of IPV,’ ‘relationship problems,’ ‘excessive jealousy or control,’ ‘employment or financial problems,’ ‘suicidal/self-harming behaviours,’ ‘antisocial attitudes,’ ‘substance use during/leading up to current IPV,’ and ‘victim fear of future IPV.’

Challenges with Coding Dynamic Variable

Unfortunately, the ‘escalation of IPV’ item was removed from consideration as a dynamic risk factor in the current study because it had poor interrater reliability between the two graduate student coders. This poor interrater reliability may have been the result of the operationalization of the variable and the shortcomings of data that were available to score the variable as intended. That is, escalation was operationalized as an increase in the frequency or severity of attempted or actual IPV, consistent with most previous research. However, other research on escalation has typically collected data over several time points to naturally observe escalations in IPV, or police officers explicitly asked victims about escalations in IPV and documented this in their file notes (Barnham et al., 2017; Bland & Ariel, 2015; Piquero et al., 2006). As a limitation of the current study, the details about escalating frequency or severity of IPV behaviours was not explicitly

recorded in file notes, and pre-index IPV files for the suspects were not available to review for signs of possible escalation. Instead, attempts were made in vain to code information about escalation in IPV prior to the index using sparse notes made by officers in the police files, which likely helps to explain the subjective and unreliable ratings of escalation in this study.

Dynamic Risk Factors and IPV Recidivism Outcomes

The remaining seven potential dynamic risk factors were examined according to Douglas and Skeem's (2005) criteria, which stipulate that for a risk factor to be dynamic, it must be correlated with the recidivism outcome and changeable over time. The dynamic risk factors examined as part of the current study were selected on the basis that they have been shown in previous research to be correlated with IPV recidivism. However, in the current study, the 'relationship problems,' 'employment/financial problems,' 'excessive jealousy/control,' 'substance use,' and 'suicidality/self-harm' items were not significantly correlated with the IPV recidivism outcome variables. Only the 'antisocial attitudes' and 'victim fear or certainty of future IPV' items were significantly correlated with at least one IPV recidivism outcome. That is, 'antisocial attitudes' was correlated with general IPV recidivism and 'victim fear' was correlated with recidivism specifically in the form of physical IPV and stalking/harassment. In addition, 'antisocial attitudes' was the only dynamic risk factor that was significantly related to time to recidivism, with increased ratings of antisocial attitudes related to faster time to recidivism.¹² The time to recidivism ranged from an

¹² Follow-up analyses indicated that the 'antisocial attitudes' variable was only significantly related to the IPV recidivism outcome variables for Caucasian suspects, but not for BIPOC suspects. This calls into question the generalizability of these findings. Future research should incorporate suspect race

average of 44 weeks for the lowest score on the ‘antisocial attitudes’ item to 41 weeks for the highest score, indicating that this item is likely tapping into stable dynamic risk (rather than imminent risk).

The finding that five of the dynamic risk variables were not significantly correlated with the IPV outcome variables does not necessarily indicate that there is a lack of validity to the constructs as dynamic risk factors for IPV. Rather, it suggests that there may have been problems with the way these constructs were measured and the quality of the data in the current study. Previous research examining these dynamic factors have used a variety of methodologies, most frequently: (a) prospective studies that have used self-report surveys or interviews to capture the suspects’ or victims’ perception of dynamic risk factors (e.g., Cattaneo & Goodman, 2003; Gondolf, 2002; Goodman et al., 2000; Heckert & Gondolf, 2002; Lamis et al., 2013; O’Keefe, 1998; Weisz et al., 2000), (b) prospective studies that collect data about dynamic risk factors in real time, which can then be used to examine imminent risk (e.g., Fals-Stewart, 2003; Moore et al., 2011; Parks et al., 2008), (c) retrospective studies that use self-report data from surveys or interviews to capture the involved parties’ perception of the presence and timing of dynamic risk factors (e.g., Campbell et al., 2003; Elisha et al., 2010; Hanson et al., 1997; Murphy et al., 2005), and (d) retrospective studies using file review of a variety of types of existing records (e.g., Belfrage & Rying, 2004; Dobash et al., 2004; Flynn et al., 2016; Grann & Wedin, 2002; Petersson et al., 2016).

demographics into the methodology and planned analyses from the inception of the research to responsibly examine how these and other findings generalize to BIPOC individuals. This is discussed further in the Conclusions section of this dissertation.

Despite the use of various methodologies in the scientific literature on dynamic risk factors, previous research has not attempted to retrospectively code dynamic factors from police file notes in the way that was done for the current study. Those previous studies that have used retrospective file coding had one of two advantages: access to files that included intentional documentation of dynamic risk factors (i.e., police documented dynamic risk factors in their file notes at the time of their investigation as part of completing a dynamic risk assessment tool; e.g., Grann & Wedin, 2002; Petersson et al., 2016), or access to multidisciplinary files with extensive documentation because the samples were comprised of intimate partner homicide cases (e.g., Belfrage & Rying, 2004; Dobash et al., 2004; Flynn et al., 2016). Because of their nature, IPV homicide files contain substantially more information and evidence compared to non-homicide police files. By comparison, the notes available in the police files accessed in the current study were typically very focused on their areas of investigation and did not include any cases on intimate partner homicide.

The challenges with the police report data in the current dissertation could be summarized as issues with data quality and quantity. Information relevant to the coding of the dynamic risk constructs of interest was not consistently or thoroughly documented, ostensibly because the responding officers were not looking for or assessing many of these dynamic risk factors. This meant that the coders in the present study were often required to attempt to infer evidence of dynamic risk from the documented impressions of the responding officers and the scripts of conversations that they recorded in their files. At other times, there was no information from which to infer, and therefore the dynamic risk factor was scored zero (i.e., as no evidence for the item).

Unfortunately, there was no way to distinguish between a dynamic risk factor that was absent or irrelevant versus information simply not being documented (i.e., missing data). To make such a distinction, the officer's file note would need to expressly document the absence of the factor (e.g., "the suspect did not demonstrate excessive jealousy or controlling tactics," or "the suspect has a stable job and source of income"). Unfortunately, officers' notes rarely documented the absence of a dynamic risk factor. This practice means that there may have been instances in which dynamic risk factors were rated as being absent when, in reality, they were present but undocumented.

The concern about the quality and quantity of the data available to code the dynamic risk factors examined in the current study does not apply equally to all seven of these variables. That is, there were three items for which there was often more information in the police file to draw from. Those items included the 'suspect substance use prior to or during IPV,' 'victim fear/certainty of future IPV, and 'suspect demonstrates antisocial attitudes.' Of these three dynamic items, the one that was the most consistently straightforward to score from police files was the 'substance use' item, presumably because documenting the suspected intoxication of involved parties is a routine part of police investigations and scoring of the ODARA. The problem with the validity of this item was more likely the result of the operationalization being too sensitive and not specific enough. The dynamic 'substance use' item was operationalized in such a way that if the suspect had any amount of alcohol or drugs prior to or during the IPV incident, then it was coded as present. Previous research on substance use as a dynamic risk factor of IPV is robust and has been replicated with a variety of different methodologies, but one thing these studies have in common is that

'substance use' was much more narrowly defined. For instance, past research has used terms such as "excessive alcohol use," "binge drinking," "alcohol problems," "increased alcohol consumption," "heavy drinking days," and "significantly higher number of drinks consumed" in the operationalization of substance use (Barnett & Fagan, 1993; Fals-Stewart, 2003; Gondolf, 2002; Pan et al., 1994; Parks et al., 2008; Murphy et al., 2005). In hindsight, the operationalization of the 'substance use' item in the current study was likely overly sensitive to have much utility. Future research should continue to examine *problematic* or *impairing* substance use as a dynamic risk factor for IPV, rather than *any* substance use.

The 'victim fear' item was another dynamic risk factor that often had richer data to draw from within the police files. The ease of coding this item may have been because it was more often documented as part of the officer's ODARA ratings, but this explanation is limited because many files did not include a police-scored ODARA. The more frequent documentation of victim fear of future IPV also may be the result of officers often noting it in the context of the victim's lack of cooperation with the investigation. As previously noted, police in New Brunswick need to collect evidence and present it to the Crown in making a request or recommendation for charges to be laid. The Crown approves all charges and attempts to ensure that only cases with a reasonable prospect of conviction and in the public interest proceed (Office of Attorney General, 2017). It was common to see police officers document the level of victim fear in their notes to the Crown, especially in cases when the victim was unwilling to make a statement as evidence for court, possibly to encourage the Crown to consider laying

charges even without the victim's cooperation and for bail considerations should the suspect be charged.

Finally, the 'antisocial attitudes' item also was often easier to code in the current study due to the presence of sufficient information in the police files. The improved quality and quantity of relevant information for this dynamic factor may be, in part, because some behaviours that could be coded as evidence of antisocial attitudes are chargeable offences, such as violations of court orders at the time of the index (e.g., IPV against a victim for whom they have a no contact order in place), resisting arrest, assaulting the police officer, or obstruction of justice. Other times, behaviours that represent antisocial attitudes were documented as a justification for removing the suspect from the scene or use of force (e.g., if the suspect was verbally berating and blaming the victim in front of the police officer, if the suspect attempted to assault the officers, or if the suspect was excessively confrontational with the officer and blocking their access to the victim or children that were present). That is to say, the behaviours that represent the 'antisocial attitudes' item appear to be more relevant to a police officer's interactions and immediate decision-making than various other dynamic risk factors (e.g., employment/financial problems, relationship problems, jealousy/control).

Although it was disappointing that the majority of potential dynamic risk factors that were explored were not correlated with IPV recidivism outcomes in the current study, the inconsistency of these findings with previous research on the same dynamic risk factors provides an important lesson for future research. That is, it illustrates the difficulty of relying on the quality of information collected in police files for such research purposes. When police officers take notes after responding to an IPV call, their

main objective is to document the information required to justify their actions taken (e.g., arrest, use of force), record evidence, and to secure certain outcomes for the suspect within the criminal justice system (e.g., charges, conditions of release). This procedure can be at odds with the information sought by researchers, who are interested in specific variables of theoretical or scientific interest that may not be routinely recorded.

Although previous research has employed other methodologies with more experimental control to investigate dynamic risk factors of IPV with more success (e.g., surveys or interviews with victims or offenders), there are some benefits to using police file review in research. For instance, police file review research is more amenable to random sampling, circumventing participant self-selection bias, and non-interference with the frontline data collection process (Alison et al., 2001). File review is also an unobtrusive method of data collection that does not require participants to actively engage in the research process. In addition, developing a risk tool based on data extracted from police reports ensures that the resulting tool will be realistically scorable by police officers using information they have access to as part of their routine investigations.

Instead of abandoning the methodology of police file review, Alison et al. (2001) recommended that researchers be involved in designing procedures for information collection that would work within the limits of police field work and training police officers in record keeping. This would allow researchers to help guide what and how information gets recorded in preparation for prospective or future research. Alternatively, researchers can be selective about the types of files reviewed in

retrospective research. To illustrate, consider the method employed by researchers such as Grann and Wedin (2004) and Petersson et al. (2016), who examined dynamic risk factors for IPV using retrospective review of police files. In both studies, police had been trained to collect and document dynamic risk information as part of their investigations (Grann & Wedin, 2004; Petersson et al., 2016). Therefore, when these researchers retrospectively coded information using police file review to examine dynamic risk factors for IPV, the relevant information was readily available in the files. Future research may be able to use such methodologies to offset the limitations of retrospective police file review, and thereby provide more useful information about dynamic risk factors for IPV.

Changeability of Dynamic Risk Factors

The next step in the current study was to examine the changeability of each dynamic risk item, which is another of Douglas and Skeem's (2005) precursors for a risk factor to be considered dynamic. The results of these analyses indicated that each of the proposed dynamic risk factors were able to reliably fluctuate over time. Of course, all this really means is that these factors can change (as per their operationalization for the current study). This change could be caused by any number of factors, such as changes in the suspect's age, job, finances, access to substances, place of residence, relationship status, childcare support, participation in therapy, police intervention, or a myriad of other possible factors. A limitation of the reliable change analyses conducted in the current study is that they are unable to inform whether or not police intervention influenced these changes, nor if change in the measured dynamic risk influenced recidivism (Heffernan et al., 2019). It is crucial that future research examines changes in

potential dynamic risk factors for IPV before and after targeted police intervention using a control group for whom these factors are not targeted. This focus of research will allow for a deeper understanding of the value of police scoring a supplemental dynamic risk tool and using it to guide their risk management decisions (Heffernan et al., 2019).

Internal Validity of the Dynamic Risk Supplemental Tool

Based on the data emerging from the current examination of dynamic risk factors, it became clear by this point that many of the dynamic risk factors (as measured in the current study) were poor candidates for a dynamic risk supplement to the ODARA. Nevertheless, the remaining step of this work was to examine the internal validity of the seven dynamic risk items as a potential dynamic scale. The item scores were summed to produce a composite score that would represent the supplemental dynamic risk assessment tool. The expectation was that, if the items were effectively measuring the underlying concepts (hypothesized to be risk for IPV), then the internal consistency analyses would find that all the items correlated with the total dynamic risk score. Given the results discussed so far, the outcome of tests of internal validity of the dynamic composite were not surprising. None of the items had a strong relationship with the total scale, although the ‘antisocial attitudes’ and ‘victim fear’ items did have a moderate relationship with the total dynamic risk score. The ‘substance use’ item, which again was likely defined too generally, had the lowest item-total correlation.

Examination of the intercorrelations between the dynamic risk factors similarly demonstrated that the only two items with acceptable multicollinearity were the ‘antisocial attitudes’ and ‘victim fear’ items. As a whole, the dynamic risk items that made up the supplemental tool showed poor internal validity, and no items were

disproportionately contributing to the problem (i.e., there were no items that, if deleted, would improve the internal consistency).

Although the original goal of the current study was to develop a supplemental tool for the ODARA that could estimate dynamic risk and help guide police decision-making, the results of the analyses discussed so far demonstrated that only the ‘antisocial attitudes’ and ‘victim fear’ items were correlated with IPV recidivism outcomes and internally valid. Thus, the other five dynamic risk factors were removed from further consideration. With two items remaining, it no longer made sense to develop a separate tool, especially considering that the ‘victim fear’ item is already an ODARA item. In essence, the ‘antisocial attitudes’ item was the only new factor to consider introducing to the ODARA.¹³ Therefore, the analyses of incremental validity were limited to examining and comparing the original ODARA to a revised version of the ODARA that included the ‘antisocial attitudes’ item (referred to as “ODARA + AA”). To be consistent with the scoring of the ODARA and to uphold the intention of the ODARA to be simple for police to fill out, the ‘antisocial attitudes’ item was converted to a dichotomous item such that a partial score was converted to 1, capturing evidence of this item.

Evaluation of the ODARA + AA

Incremental Predictive Validity of ODARA + AA

The ODARA was moderately predictive of IPV recidivism at the one-, three-, six-, and twelve-month mark, whereas the ‘antisocial attitudes’ item was not predictive

¹³ As discussed in footnotes throughout this dissertation, follow-up analyses with the ‘antisocial attitudes’ item suggest that there may be inherent bias towards BIPOC individuals with the operationalization of this item. This finding calls into question the appropriateness of this item, which is discussed further in an addendum that follows the “Discussion” section.

of above chance at each of these time points. However, there is little value to examining the ‘antisocial attitudes’ item as a unitary concept because it is not expected that any crime or behaviour can be explained or predicted by one factor alone. Bonta and Andrews (2017) proposed that dynamic risk factors accumulate and interact with other factors to influence behaviour; therefore, the ODARA was compared to the ODARA + AA in terms of predictive validity. It was hypothesized that the addition of the ‘antisocial attitudes’ item would provide incremental validity at the one- and three-month periods, based on the findings of comparable previous research (Chu et al., 2013; Finch et al., 2016). The results indicated that ODARA + AA was moderately predictive of IPV recidivism at each of the four time points. Although the addition of the ‘antisocial attitudes’ item did not provide statistically significantly incremental validity at any of the time points, the effect sizes did increase at each time point. The ODARA + AA was also examined for incremental validity in its ability to predict a suspect’s time to first recidivism above and beyond the ODARA within the 12-month follow-up period. This analysis found that, although the ODARA + AA could significantly predict time to recidivism, it did not provide statistically significant incremental validity over the ODARA on its own. This finding could stem from the fact that there may be items on the ODARA that indirectly or directly tap into antisocial values or attitudes given their overlapping criminal content.

Changeability of ODARA + AA

The final evaluation of the ODARA + AA was to examine its ability to fluctuate over time and compare it to the original ODARA. It was hypothesized that the addition of a dynamic risk factor would improve the sensitivity of the tool to detect change in risk

over time. The results indicated that the ODARA + AA was, in fact, marginally yet significantly more sensitive to fluctuations of risk over time. Suspects' scores on both the ODARA and ODARA + AA in general were more likely to increase between the index and the first recidivism event. This pattern makes sense given the fact that the subsample for this analysis was, by definition, limited to recidivists only (i.e., there needed to be at least two police contacts in order to examine change in suspects' risk scores). However, when examined at the level of risk classification (low-, moderate-, and high-risk suspects), the results demonstrated that while the highest risk suspects were most likely to have an *increase* or no change in their original ODARA risk score, they were most likely to have no change or a *decrease* in risk in their ODARA + AA risk score. This nuance indicates that the simple addition of the 'antisocial attitudes' item improves the tool's ability to detect both increases and decreases in risk.

Implications and Conclusions

Implications for Police Use of the ODARA

The results of the current study demonstrated that the ODARA lacked predictive validity when used by frontline police officers in New Brunswick. More specifically, the results indicate that there is a need to improve police compliance with IPV risk assessment policy and scoring protocol of the ODARA, which in turn could improve the utility of the tool to guide police decision-making and mediate recidivism. Two particular findings suggest that at least part of the issue in implementing the ODARA in this sample stemmed from a lack of appropriate use by frontline officers: (1) about half of the police-scored ODARAs were not completed in full (i.e., had missing data, totals summed inaccurately, proration incomplete), and (2) some of the items with the worst

reliability were the very items that could be scored by accessing the suspect's police record, which is readily available to police. Although the current study did not formally examine police feedback on the ODARA and its implementation, these findings may point to a lack of individual buy-in to the ODARA, social pressures to comply with related policy, and perception of control and understanding of the risk assessment process.

Although an in-depth examination of gender issues with IPV risk assessment was beyond the scope of the current dissertation, preliminary analyses offered an important reminder that the issue of gender remains critically important to consider. For instance, the current study found the ODARA performed similarly for male and female suspects when relying on the scores from researchers' retrospective coding of the police files, but there were notable differences in how police used the ODARA with male and female suspects. That is, police officers in this sample were more likely to complete the ODARA with male suspects, to incorrectly classify female suspects as low-risk and male suspects as high-risk, to over-intervene with male suspects, and to offer more discretionary aid to both female victims and female suspects. These findings demonstrate that even though actuarial risk assessment tools such as the ODARA can be used to improve risk estimation above professional judgement, it does not eliminate the possibility of influences from biases and errors in human judgment (Ægisdottir et al., 2006; Bloom et al., 2005; Dawes, 1979; Litwack & Schlesinger, 1999; Northcott, 2012; Quinsey et al., 1998; Roehl & Guertin, 2000). These results call for research to continue to closely examine the sources and impact of such biases and improve police training and oversight to improve risk assessment scoring, validity, and utility.

In the best-case scenario, an unreliably scored risk assessment tool is a waste of time and money (Dvoskin et al., 2012), both of which are limited resources in policing. A far worse-case scenario is that an incorrectly scored risk assessment tool can result in over-intervening with low-risk offenders and under-intervening with high-risk offenders, which may increase the likelihood of recidivism (Belfrage et al., 2011; Dvoskin et al., 2012; Nicholls et al., 2013; Williams & Stansfield, 2017). However, the worst-case scenario is that individuals who are high-risk for violent recidivism are misclassified and there are insufficient risk mitigation strategies put in place (e.g., no detention, release on bail, no follow-up checks or monitoring, no conditions) and suspects have increased opportunity to revictimize their target. For all these reasons, it is worth increasing efforts to improve risk assessment training, buy-in, and supervision. Future research should examine police compliance with IPV risk assessment protocols, accuracy of field scoring, predictive validity, application to guide risk management decision-making, and effect on recidivism before and after the changes to education, training, and supervision are implemented. In addition, these outcomes should be assessed at multiple time points after these changes are made to further understand how time and other factors might affect procedural drift, and to inform how frequently retraining is required.

An additional area for future research with regards to police field use of the ODARA involves the possibility of developing a secure, web-based app added to their mobile devices or vehicle console. The ODARA is easily amenable to automation, and an app-based program could be used to prompt officers to fill out missing data and could provide a quick reference to item definitions while they are completing the tool. In fact, the possibilities are vast when considering combining the power of risk assessment,

scientific data, and technology. For instance, if integrated with a secure data storage system, an app would allow for many of the historical/static risk items of the ODARA to be filled automatically based on file information (without the need to risk civilians listening on police-scanners when communicating suspect or victim information over the radio with dispatch to collect file information in the field). These technological possibilities could help reduce errors in scoring and missing data. In addition, an automated electronic version of the ODARA could allow for the presence of certain items or the final risk level information to be flagged by the app to prompt the officer to consider specific risk management strategies (e.g., if the suspect confined the victim at the index, they may consider holding the suspect overnight, issuing a no-contact order, and/or helping the victim find safe accommodations). In fact, this could create many further opportunities for the future of RNR research in terms of examining the nuances of risk-need-responsivity match (e.g., considering gender, race, sexual orientation/identification, living arrangement, etc.). It is important that future research examines which specific risk management strategies are most effective in reducing recidivism for suspects posing various levels of risk and with particular criminogenic needs, in keeping with the RNR model.

Implications Regarding Dynamic Risk Factors for IPV

In the end, there was only one novel item that was examined for incremental validity above and beyond the ODARA. The ‘antisocial attitudes’ item did not add to the predictive validity in a statistically significant way; however, it could offer practical value. More specifically, the ‘antisocial attitudes’ item did improve the tool’s ability to detect changes in risk. In theory, this added sensitivity could help police detect

fluctuations in risk over time or across situations. This information could be used to detect escalations in risk that may lead to preventable violent outcomes. It also could increase the ODARA's utility for police to examine the effectiveness of their responses and risk management decisions. That is, a more dynamic risk tool that can detect both increases and decreases in risk could be used to observe if police intervention has any impact on fluctuations in a suspect's risk over time.

Another possible benefit of including the 'antisocial attitudes' item in the ODARA is that it could help guide police decision-making, as would the addition of any dynamic risk factor. For instance, past research has demonstrated that a suspect with evidence of antisocial attitudes is more likely to breach court orders (Eastman et al., 2019). Therefore, if a high-risk suspect is demonstrating antisocial attitudes, then this information could help inform the responding officer's decision to make an arrest and recommend to the Crown that the suspect is held until trial (rather than recommending a conditional release, since they are more likely to violate those conditions).

Despite the possibility of increased practical utility, the overall results of this dissertation do not warrant altering the ODARA to add the 'antisocial attitudes' item at this time, given its methodological limitations. However, to entirely dismiss the 'antisocial attitudes' item – or other dynamic risk factors – based on the results of this study could also be an unfortunate oversight. Rather, the results of this dissertation emphasize the need for careful methodological planning when exploring dynamic risk factors for IPV (such as prospective, longitudinal, multi-site studies with large and diverse samples; Klepfisz et al., 2016). If future research addresses the methodological limitations of the current study, it has the potential to enhance the ODARA's ability to

capture changes in risk over time, prioritize imminent risk mitigation actions, and target longer-term risk management strategies by non-police intervention supports to reduce the impact of these factors. Considering the ODARA is used by police services across New Brunswick and various other provinces in Canada, such research could have practical implications that are far-reaching.

Limitations

Various limitations of the current study and possible implications for future research have been discussed throughout this dissertation and are not reiterated here. However, there are some notable limitations that remain to be discussed about using police file review for data collection. One such limitation involves sampling. Although random selection was used as much as possible to limit systematic bias (with the exception of some quota-based random sampling to ensure enough female suspects were included in the sample), there may have been other systematic factors outside of research control that introduced bias to the sample. That is, thousands of IPV police files were reviewed in order to identify the 309 index files that included sufficient information to be included in the final sample. Many of the IPV files reviewed had only a few sentences to document the event. Unfortunately, it is not known if there were any patterns in which types of IPV scenarios were less likely to be thoroughly documented. Presumably the non-chargeable IPV acts would have less documentation (e.g., emotional abuse such as name calling, gaslighting, controlling the victim's social life, hiding the car keys, spreading rumors so the victim's friendships are jeopardized, etc.); however, people also may be less likely to call the police in the first place when the IPV does not include a chargeable act. It is also possible that IPV events in which the victim is

unwilling to give a statement and no other evidence is available (e.g., no visible injury, no witnesses, no suspect statement) have less documentation as the officer might have no intention of recommending charges without being able to present evidence to the Crown.

An additional limitation of police file review research is incompleteness of the recidivism data. A lack of IPV-related contact with the police certainly does not equate to a lack of IPV recidivism. First, a great deal of IPV goes unreported to the police and, therefore, the rates of recidivism found in this study are almost certainly an underrepresentation of the true recidivism rates (Burczycka, 2018). Second, it is possible that individuals recidivated in a different jurisdiction and, therefore, the event would not appear in their police file reviewed for the current study. Third, it is possible that failure to recidivate in the 12-month follow-up period was due to the suspect being held in police custody until trial, convicted, and incarcerated and, therefore, had no opportunity to recidivate. Pre-trial holdings and incarcerations are often, but not always documented in police files. When this information was available, these files were screened out of relevant analyses. Finally, it is possible that some suspects deceased prior to the end of the 12-month follow-up period, which may not be documented in police files.

Imperfection in the measurement of offending and reoffending can affect all research on violent or criminal behaviour, requiring realistic expectations for research on risk assessment (Harris & Rice, 2003). However, even though much of IPV goes unreported to the police and as such is not captured in studies such as the current dissertation, available research has demonstrated that those individuals who commit the majority of unreported offenses are the same people who commit most of the detected

offenses (Farrington, 2002). In addition, previous research has found that the patterns are similar for predicting police- and self-reported criminal or violent behaviour (Farrington, 1985; Monahan et al., 2001; Murphy et al., 2003). These findings suggest that a suspect's risk ranking (i.e., relative to other suspects) can be expected to be similar for official or self-report data (Mental Health Centre Penetanguishene Research Department & Ontario Provincial Police Behavioural Sciences Section, 2005). In the end, the use of police reports for data collection was the best method available for the current dissertation, but it is an imperfect source of data. Therefore, it should be assumed that the current study has underestimated the actual rates of reoffending. In addition, the ODARA scores produced for suspects in this study should be limited to predicting the likelihood of a future IPV-related police contact (which is consistent with the intent of the ODARA; Hilton et al., 2004).

An Addendum: Considerations about Race and Future Research

In 2020, when police actions led to the tragic deaths of George Floyd and Breonna Taylor in the United States and Rodney Levi-Peters, Chantel Moore, and Regis Korchinski-Paquet in Canada, a renewed international outcry of indignation took place against systemic racism and racialized police violence. These events have shone a spotlight on the fact that systemic changes are needed to address the issue of racism, and this is a responsibility that must be upheld within the field of psychology and psychological research. The fact that the effects of suspect race were very nearly overlooked in the current dissertation is an illustration of this systemic racial oppression, especially given the subject of policing in this research. Regrettably, I did not identify this oversight until doing one of the final read-throughs in the last few days before my

dissertation submission. Although this meant there were limitations of feasibility in terms of the extent to which I could address issues of race in this dissertation, I recognized that there was an ethical imperative minimally to draw attention to this negligence and to implore future researchers to learn from this mistake.

Failing to examine the role of suspect race in police research might lead to important information being missed or misrepresented. For instance, I included behaviours that indicate that the suspect was opposed to police presence or involvement as part of the operationalization of the “antisocial attitudes” risk factor in this dissertation. It was only after my dissertation was completed that I recognized how this could potentially misrepresent BIPOC suspects, who may demonstrate opposition to police presence or involvement for reasons that have nothing to do with having antisocial attitudes. Applying an “anti-racist” lens to this variable can provide a clearer understanding of why the operationalization is problematic. The reality is that throughout Canadian history, BIPOC individuals have been more likely to experience violence at the hands of public safety officials, including increased susceptibility to police and correctional officer use of force, violence, and killings (Comack, 2012; Maynard, 2017; Sharpe, 2016). For instance, a literature review published in 2006 found that, despite making up three percent of Canada’s population at the time, Black Canadians made up about 30% of those who are killed by police (Wortley, 2006). Other research has found that police use of force is five times more severe for Black Canadians compared to White Canadians (Sharpe, 2016). As Black Studies scholar Christina Sharpe (2016) puts it, to be Black in Canada is not just to be over-policed, it is to be “proximate to death” during police encounters (p. 16). Likewise, Indigenous Canadians

have experienced disproportionate systemic police violence in the forms of use of force, police shootings, “star light tours” (i.e., the often lethal police practice of dropping off Indigenous people in remote areas in the freezing cold, forcing them to find their way home), and “phone bookings” (i.e., the police practice of holding a phonebook to the head or body of an Indigenous person prior to beating them, in order to minimize evidence in the form of bruising; Cormack, 2012). The most recent national data available suggests that BIPOC individuals make up 51% of individuals involved in public safety officers’ use of force cases since 2015, while representing only 8.5% of the Canadian population (Zinger, 2021).

Given the historical and current treatment of BIPOC individuals within the Canadian criminal justice system, it is not surprising that negative attitudes about the police persist in various BIPOC communities (Barrett & Peirone, 2019). As such, it is a strong possibility that various behaviours of opposition to police presence during IPV-related encounters represent a fear of the police, anger towards the police, or distrust of the police (rather than holding true antisocial attitudes). In fact, the preliminary follow-up analyses suggested that the ‘antisocial attitudes’ variable did not predict IPV recidivism at 1, 3, 6, or 12 months, correlate with frequency of IPV recidivism, or predict quickness to IPV recidivism with BIPOC suspects, whereas it did for Caucasian suspects.

Although the ‘antisocial attitudes’ item in the current study did not appear to be relevant for BIPOC individuals, it does not necessarily mean that *true* antisocial attitudes are an unimportant risk factor for this group. The true effect of the ‘antisocial attitudes’ risk factor might be washed out with BIPOC individuals in the current sample because

of how the variable was operationalized (i.e., because of the racially-biased inclusion of anti-authority behaviours). Unfortunately, these effects could not be teased apart because of the way the data were coded in this study. The interpretation of these findings is limited due to the small number of BIPOC individuals in this sample, but it does emphasize the necessity for researchers to consider issues of race in planning research methodologies and to statistically examine the effects of race. It also reaffirms my previous conclusion that future research should address the methodological limitations of the current dissertation and continue to explore dynamic risk factors for IPV, with more attention to the many ways that racial bias could infiltrate research and influence the interpretation of the results.

References

- Ægisdottir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., Nichols, C. N., Lampropoulos, G. K., Walker, B. S., Cohen, G., & Rush, J. D. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counselling Psychologist, 34*(3), 341-382.
<https://doi.org/10.1177/0011000005285875>
- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage Publications.
- Ajzen, I. (1991). The theory of planned behavior. *Organization Behavior and Human Decision Processes, 50*(2), 179-211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Alpert, G. P., Dunham, R. G., Stroshine, M., Bennett, K., & MacDonald, J. (2006). Police officers' decision making and discretion: Forming suspicion and making a stop. *Department of Justice*. <https://www.ojp.gov/pdffiles1/nij/grants/213004.pdf>
- Alison, L., Snook, B., & Stein, K. L. (2001). Unobtrusive measurement: Using police information for forensic research. *Qualitative Research, 1*(2), 241- 254.
<https://doi.org/10.1177/146879410100100208>
- American Bar Association. (1980). *Standards for criminal justice* (2nd ed., Vol. 1). Boston, MA: Little, Brown.
- Amnesty International (2014). Violence against indigenous women and girls in Canada: A summary of Amnesty International's concerns and call to action. *Amnesty*

International Canada. https://www.amnesty.ca/sites/amnesty/files/iwfa_submission_amnesty_international_february_2014_-_final.pdf.

Andrews, D. A. (2015). *Recidivism is predictable and can be influenced: Using risk assessments to reduce recidivism*. Retrieved from Correctional Service Canada website: https://www.csc-scc.gc.ca/research/forum/special/espe_a-eng.shtml

Andrews, D. A., & Bonta, J. (2003). *The psychology of criminal conduct* (3rd ed.). Cincinnati, OH: Anderson Publishing.

Andrews, D. A., & Bonta, J. (2006). *The psychology of criminal conduct* (4th ed.). Newark, NJ: Lexis Nexis.

Andrews, D. A., & Bonta, J. (2010). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and Law*, 16(1), 39–55.
<https://doi.org/10.1037/a0018362>

Andrews, D. A., Bonta, J., & Wormith, J. S. (2004). *Level of Service/Case Management Inventory: An offender assessment system*. North Tonawanda, NY: Multi-Health Systems.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The recent past and near future of risk and/or need assessment. *Crime and Delinquency*, 52(1), 7-27.
<https://doi.org/10.1177/0011128705281756>

Andrews, D. A., Bonta, J., & Wormith, J. (2011). The Risk-Need-Responsivity (RNR) model. *Criminal Justice and Behavior*, 38(7), 735-755.
<https://doi.org/10.1177/0093854811406356>

Andrews, D. A., Zinger, I., Hoge, R., Bonta, J., Gendreau, P., & Cullen, F. (1990). Does correctional treatment work? A clinically-relevant and psychologically-informed

meta-analysis. *Criminology*, 28(3), 369-404. https://www.researchgate.net/profile/James_Bonta2/publication/264637995_Does_Correctional_Treatment_Work_A_Clinically_Relevant_and_Psychologically_Informed_Meta-Analysis/links/5b94140a299bf147392a172d/Does-Correctional-Treatment-Work-A-Clinically-Relevant-and-Psychologically-Informed-Meta-Analysis.pdf

Ansara, D. L., & Hindin, M. J. (2010). Formal and informal help-seeking associated with women's and men's experiences of intimate partner violence in Canada. *Social Science & Medicine*, 70(7), 1011–1018.

<https://doi.org/10.1016/j.socscimed.2009.12.009>

Au, A., Cheung, G., Kropp, R., Yuk-chung, C., Lam, G. L. T., & Sung, P. (2008). A preliminary validation of the Brief Spousal Assault Form for the Evaluation of Risk (B-SAFER) in Hong Kong. *Journal of Family Violence*, 23(8), 727-735.

<https://doi.org/10.1007/s10896-008-9198-z>

Avakame, E. F., & Fyfe J. J. (2001). Differential police treatment of male-on-female spousal violence: Additional evidence on the leniency thesis. *Violence against Women*, 7(1), 22–45. <https://doi.org/10.1177/10778010122182280>

Ballucci, D., Gill, C., & Campbell, M. A. (2017) The power of attitude: The role of police culture and receptivity of risk assessment tools in IPV calls. *Policing: A Journal of Policy and Practice*, 11(3), 242–257.

<https://doi.org/10.1093/policing/pax018>

Ballucci, D., Campbell, M. A., & Gill, C. (2020). Use of the ODARA by police officers for intimate partner violence: Implications for practice in the field. *Journal of*

Community Safety and Well-Being, 5(3), 91-102.

<https://doi.org/10.35502/jcswb.150>

Barnett, G. D., Wakeling, H., Mandeville-Norden, R., & Rakestrow, J. (2013). Does change in psychometric test scores tell us anything about risk of reconviction in sexual offenders? *Psychology, Crime and Law*, 19(1), 85-110.

<https://doi.org/10.1080/1068316X.2011.607820>

Barnett, O. W., & Fagan, R. W. (1993). Alcohol use in male spouse abusers and their female partners. *Journal of Family Violence*, 8(1), 1-1.

<https://doi.org/10.1007/BF00986990>

Barnham, L., Barnes, G. C., & Sherman, L. W. (2017). Targeting escalation of IPV: Evidence from 52,000 offenders. *Cambridge Journal of Evidence-Based Policing*, 1, 116-142. <https://doi.org/10.1007/s41887-017-0008-9>

Barnes, P. G. (1998). It's just a quarrel. *American Bar Association Journal*, 84(2), 24-25.

Barrett, B. J., & Peirone, A. (2019). Spousal violence and evaluations of police performance in Canada: Does police contact matter? *Journal of Family Violence*, 34(3), 199-211. <https://doi.org/10.1007/s10896-018-0017-x>

Bartko, J. J., & Carpenter, W. T. (1976). On the methods and theory of reliability. *Journal of Nervous & Mental Disease*, 163(5), 307-317.

<https://doi.org/10.1097/00005053-197611000-00003>

Barton, H., 2003. Understanding occupational (sub) culture-a precursor for reform. *International Journal of Public Sector Management*, 16 (5), 346-358.

Beggs, S. M., & Grace, R. C. (2010). Assessment of dynamic risk factors: An independent validation study of the Violence Risk Scale: Sexual Offender

version. *Sexual Abuse: A Journal of Research and Treatment*, 22(2), 234–251.

<https://doi.org/10.1177/1079063210369014>

Beggs, S. M., & Grace, R. C. (2011). Treatment gain for sexual offenders against children predicts reduced recidivism: A comparative validity study. *Journal of Consulting and Clinical Psychology*, 79(2), 182-92.

<https://doi.org/10.1037/a0022900>

Belfrage, H., & Rying, M. (2004). Characteristics of spousal homicide perpetrators: A descriptive study of all cases of spousal homicide in Sweden 1990-1999.

Criminal Behaviour and Mental Health, 14(2), 121-133.

<https://doi.org/10.1002/cbm.577>

Belfrage, H., & Strand, S. (2001). Comparison of HCR-20 scores in violent mentally disordered men and women: Gender differences and similarities. *Psychology, Crime & Law*, 7(1), 71-79. <https://doi.org/10.1080/10683160108401784>

Belfrage, H., & Strand, S. (2008). Structured spousal violence risk assessment: Combining risk factors and victim vulnerability factors. *International Journal of Forensic Mental Health*, 7(1), 39–46. <https://www.researchgate.net/deref/http%3A%2F%2Fdx.doi.org%2F10.1080%2F14999013.2008.9914402>

Belfrage, H., & Strand, S. (2012). Measuring the outcome of structured spousal violence risk assessments using the B-SAFER: Risk in relation to recidivism and intervention. *Behavioral Sciences & the Law*, 30(4), 420-430.

<https://doi.org/10.1002/bsl.2019>

Belfrage, H., Strand, S., Storey, J. E., Gibas, A. L., Kropp, P. R., & Hart, S. D. (2011). Assessment and management of risk for intimate partner violence by police

- officers using the Spousal Assault Risk Assessment Guide. *Law and Human Behavior*, 36(1), 60–67. <https://doi.org/10.1037/h0093948>
- Belknap, J. (2001). *The invisible woman: Gender, crime, and justice* (2nd ed.). Belmont, CA: Wadsworth.
- Belknap, J. & Hartman, J. L. (2000). Police responses to woman battering: Victim advocates' reports. *International Review of Victimology*, 7(1-3), 159-177. <https://doi.org/10.1177/026975800000700309>
- Berk, R. A., Campbell, A., Klap, R., & Western, B. (1992). A Bayesian analysis of the Colorado Springs spouse abuse experiment. *Journal of Criminal Law and Criminology*, 83(1), 170-200. <https://pdfs.semanticscholar.org/5f16/21a58dfbecd8251a82f45d3f50ea10256fb.pdf>
- Bloom, H., Webster, C., Hucker, S., & Freitas, K. (2005). The Canadian contribution to violence risk assessment: History and implications for current psychiatric practice. *Canadian Journal of Psychiatry*, 50(1), 3-11. <https://doi.org/10.1177/070674370505000103>
- Bland, M., & Ariel, B. (2015). Targeting escalation in reported domestic abuse: Evidence from 36,000 callouts. *International Criminal Justice Review*, 25(1), 30–53. <https://doi.org/10.1177/1057567715574382>
- Blosnich, J., & Bossarte, R. (2012). Drivers of disparity: Differences in socially based risk factors of self-injurious and suicidal behaviors among sexual minority college students. *Journal of American College Health*, 60(2), 141-149. <https://doi.org/10.1080/07448481.2011.623332>

- Bohmer, C., Brandt, J., Bronson, D., & Hartnett, H. (2002). Domestic violence law reforms: Reactions from the trenches. *Journal of Sociology and Social Welfare*, 29(3), 71–87. <https://scholarworks.wmich.edu/jssw/vol29/iss3/5>
- Bonta, J. (2002). Offender risk assessment: Guidelines for selection and use. *Criminal Justice and Behaviour*, 29(4), 355-379.
<https://doi.org/10.1177/0093854802029004002>
- Bonta, J., & Andrews, D. A. (2007). *Risk-need-responsivity model for offender assessment and rehabilitation*. Ottawa: Public Safety Canada.
- Bonta, J., & Andrews, D. A. (2017). *The psychology of criminal conduct* (6th ed.). New York, NY: Anderson Publishing.
- Bonta, J., & Motiuk, L. L. (1990). Classification to halfway houses: A quasi-experimental evaluation. *Criminology*, 28(3), 497-506.
<https://doi.org/10.1111/j.1745-9125.1990.tb01336.x>
- Borum, R. (1996). Improving the clinical practice of violence risk assessment: Technology, guidelines and training. *American Psychologist*, 51(9), 945-956.
http://scholarcommons.usf.edu/cgi/viewcontent.cgi?article=1559&context=mhlp_facpub
- Bourgon, G., Mugford, R., Hanson, R. K., & Coligado, M. (2018). Offender risk assessment practices vary across Canada. *Canadian Journal of Criminology and Criminal Justice*, 60(2), 167–205. <https://doi.org/10.3138/cjccj.2016-0024>
- Box, G. E .P. & Tidwell, P. W. (1962). Transformation of independent variables. *Technometrics*, 4, 531-550. <https://doi.org/10.1080/00401706.1962.10490038>

- Brooks-Holliday, S., Heilbrun, K., & Fretz, R. (2012). Examining improvements in criminogenic needs: The risk reduction potential of a structured re-entry program. *Behavioral Sciences & the Law*, 30(4), 431–447.
<https://doi.org/10.1002/bsl.2016>
- Brownridge, D. A. (2008). Understanding the elevated risk of partner violence against Aboriginal women: A comparison of two nationally representative surveys of Canada. *Journal of Family Violence*, 23(5), 353–367.
<https://doi.org/10.1007/s10896-008-9160-0>
- Brunero, S., & Lamont, S. (2009). Risk analysis: An integrated approach to the assessment and management of aggression/violence in mental health. *Journal of Psychiatric Intensive Care*, 5(1), 25-32.
<https://doi.org/10.1017/S1742646408001349>
- Buchanan, K. (2009). *Risk assessment and spousal violence: Predictive validity and cultural applicability*. [Unpublished doctoral dissertation]. University of Regina.
- Burczycka, M. (2016). Trends in self-reported spousal violence in Canada, 2014. *Juristat, Statistics Canada*. (Catalogue no. 85-002-X).
<https://www150.statcan.gc.ca/n1/pub/85-002-x/2016001/article/14303/01-eng.htm>
- Burczycka, M. (2018). Section 2: Police-reported intimate partner violence in Canada, 2018. *Juristat, Statistics Canada*. (Catalogue no. 85-002-X). <https://www150.statcan.gc.ca/n1/pub/85-002-x/2019001/article/00018/02-eng.htm>

- Burke, T. W., Jordan, M. L., & Owen, S. S. (2002). A cross-national comparison of gay and lesbian domestic violence. *Journal of Contemporary Criminal Justice, 18*(3), 231–257. <https://doi.org/10.1177/1043986202018003003>
- Buzawa, E. S. (1982). Police officer response to domestic violence legislation in Michigan. *Journal of Police Science and Administration, 10*, 415-424.
- Cale, J., Tzoumakis, S., Leclerc, B., & Breckenridge, J. (2017). Patterns of intimate partner violence victimization among Australia and New Zealand female university students: An initial examination of child maltreatment and self-reported depressive symptoms across profiles. *Australian & New Zealand Journal of Criminology, 50*(4), 582-601.
<https://doi.org/10.1177%2F0004865816666615>
- Campbell, J. C. (1986). Nursing assessment for risk of homicide with battered women. *Advances in Nursing Science, 8*(4), 36-51.
<https://doi.org/10.1097/00012272-198607000-00006>.
- Campbell, J. C. (2012). Risk factors for intimate partner homicide: The importance of Margo Wilson's foundational research. *Homicide Studies, 16*(3), 438-444.
<https://doi.org/10.1177/1088767912463208>
- Campbell, J. C., & Glass, N. (2009). Safety planning, danger, and lethality assessment. In C. E. Mitchell & D. Anglin (Eds.), *Intimate partner violence: A health-based perspective*. Oxford, UK: Oxford University Press.
- Campbell, J. C., & Messing, J. T. (2017). *Assessing dangerousness: Domestic violence offenders and child abusers*. New York, NY: Springer Publishing.

- Campbell, J. C., Webster, D. W., & Glass, N. (2009). The Danger Assessment: Validation of a lethality risk assessment instrument for intimate partner femicide. *Journal of Interpersonal Violence, 24*(4), 653-674.
<https://doi.org/10.1177/0886260508317180>
- Campbell, J. C., Webster, D., Koziol-McLain, J., Block, C., Campbell, D., Curry, M. A., Gary, F., Glass, N., McFarlane, J., Sachs, C., Sharps, P., Ulrich, Y., Wilt, S. A., Manganello, J., Xu, X., Schollenberger, J., Frye, V., & Laughon, K. (2003). Risk factors for femicide in abusive relationships: Results from a multisite case control study. *American Journal of Public Health, 93*(7), 1089–1097.
<https://dx.doi.org/10.2105%2Fajph.93.7.1089>
- Campbell, M. A., French, S., & Gendreau, P. (2009). The prediction of violence in adult offenders: A meta-analytic comparison of instruments and methods of assessment. *Criminal Justice and Behavior, 36*(6), 567-590.
<https://doi.org/10.1177/0093854809333610>
- Campbell, M. A., Gill, C., & Ballucci, D. (2018). Informing police response to intimate partner violence: Predictors of perceived usefulness of risk assessment screening. *Journal of Community Safety and Wellbeing, 5*(3), 91-102.
<https://doi.org/10.35502/jcswb.150>
- Canadian Association of Chiefs of Police [CACCP]. (2016). National framework for collaborative police action on intimate partner violence (IPV). *University of New Brunswick & Canadian Associations of Chiefs of Police*.
https://cacp.ca/index.html?asst_id=1200

- Capaldi, D. M., Knoble, N. B., Shortt, J. W., & Kim, H. K. (2012). A systematic review of risk factors for intimate partner violence. *Partner Abuse, 3*(2), 231-280.
<https://doi.org/10.1891/1946-6560.3.2.231>
- Cattaneo, L. B., & Goodman, L. A. (2003). Victim-reported risk factors for continued abusive behavior: Assessing the dangerousness of arrested batterers. *Journal of Community Psychology, 31*(4), 349-369. <https://doi.org/10.1002/jcop.10056>
- Chu, C. M., Thomas, S. D., Daffern, M., & Ogloff, J. R. P. (2013). The short- to medium-term predictive accuracy of static and dynamic risk assessment measures in a secure forensic hospital. *Assessment, 20*(2), 230–241.
<https://doi.org/10.1177/1073191111418298>
- Clark, L. A. & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309-319.
<https://psycnet.apa.org/doi/10.1037/1040-3590.7.3.309>
- Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- Cohen, T. H., Lowenkamp, C. T., & VanBenschoten, S. W. (2016). Does change in risk matter? Examining whether changes in offender risk characteristics influence recidivism outcomes. *Criminology & Public Policy, 15*(2), 263-296.
<https://doi.org/10.1111/1745-9133.12190>
- Coker, A. L., Davis, K. E., Arias, I., Desai, S., Sanderson, M., Brandt H. M., & Smith, P. H. (2002). Physical and mental health effects of intimate partner violence for men and women. *American Journal of Preventive Medicine, 23*(4), 260-268.
[https://doi.org/10.1016/s0749-3797\(02\)00514-7](https://doi.org/10.1016/s0749-3797(02)00514-7)

- Coker, A. L., Smith, P. H., Thompson, M. P., McKeown, R. E., Bethea, L., & Davis, K. E. (2002). Social support protects against the negative effects of partner violence on mental health. *Journal of Women's Health and Gender-Based Medicine*, *11*(5), 465-476. <https://doi.org/10.1089/15246090260137644>
- Cole, G. F., & Gertz, M. G. (1998). *The criminal justice system: Politics and policies* (7th ed.). Belmont, CA: West/Wadsworth.
- Comack, E. (2012). *Racialized policing: Aboriginal people's encounters with the police*. Fernwood Pub.
- Conroy, S. (2021). Family violence in Canada: A statistical profile, 2019. *Juristat, Statistics Canada*. (Catalogue no. 85-002-X).
<https://www150.statcan.gc.ca/n1/pub/85-002-x/2021001/article/00001-eng.htm>
- Cooke, D. J. (1996). Predicting offending in prison: The predictive validity of the Prison Behaviour Rating Scales. *Legal and Criminological Psychology*, *1*(1), 65-82.
<https://doi.org/10.1111/j.2044-8333.1996.tb00307.x>
- Cording, J. R., Beggs Christopherson, S. M., & Grace, R. C. (2016). Challenges for the theory and application of dynamic risk factors. *Psychology, Crime & Law*, *22*(2), 84-103. <https://doi.org/10.1080/1068316X.2015.1111367>
- Cotter, A. (2018). Violent victimization of women with disabilities in Canada, 2014. *Juristat, Statistics Canada*. (Catalogue no. 85-002-X).
<https://www150.statcan.gc.ca/n1/pub/85-002-x/2018001/article/54910-eng.htm>
- Cotter, A. (2021). Intimate partner violence in Canada, 2018: An overview. *Juristat, Statistics Canada*. (Catalogue no. 85-002-X).
<https://www150.statcan.gc.ca/n1/pub/85-002-x/2021001/article/00003-eng.htm>

- Cotter, A. & Savage, L. (2019). Gender-based violence and unwanted sexual behaviour in Canada, 2018: Initial findings from the Survey of Safety in Public and Private Space. *Juristat, Statistics Canada*. (Catalogue no. 85-002-X).
<https://www150.statcan.gc.ca/n1/pub/85-002-x/2019001/article/00017-eng.htm>
- Craissati, J., & Beech, A. (2003). A review of dynamic variables and their relationship to risk prediction in sex offenders. *Journal of Sexual Aggression, 9*(1), 41–55.
<https://doi.org/10.1080/355260031000137968>
- Crenshaw, K. W. (1994). Mapping the margins: Intersectionality, identity politics, and violence against women of color. In M. A. Fineman & R. Mykitiuk (Eds.), *The public nature of private violence* (pp. 3-118). New York, NY: Routledge.
- Criminal Code*, R.S.C. 1985, c C-46, s 16(1).
- Criminal Code*, R.S.C. 1985, c C-46, s 271(a).
- Cristobal, E., Flavián, C., & Guinaliú, M. (2007). Perceived e-service quality (PeSQ). *Managing Service Quality, 17*(3), 317-340. <https://doi.org/10.1108/09604520710744326>
- Dal Grande, E., Hickling, J., Taylor, A., & Woollacott, T. (2003). Domestic violence in South Australia: A population survey of males and females. *Australian and New Zealand Journal of Public Health, 27*(5), 543-550.
<https://doi.org/10.1111/j.1467-842x.2003.tb00831.x>
- Daly, J. E., & Pelowski, S. (2000). Predictors of dropout among men who batter: A review of studies with implications for research and practice. *Violence and Victims, 15*(2), 137-160. <https://doi.org/10.1891/0886-6708.15.2.137>

- Dasgupta, S. D. (2003). *Safety & justice for all: Examining the relationship between the women's anti-violence movement and the criminal legal system*. MS Foundation for Women. http://www.ncdsv.org/images/Ms_SafetyJusticeForAll_2003.pdf
- Daoud, N., Smylie, J., Urquia, M., Allan, B., & O'Campo, P. (2013). The contribution of socio-economic position to the excesses of violence and intimate partner violence among Aboriginal versus non-Aboriginal women in Canada. *Canadian Journal of Public Health, 104*, 278–283. <https://doi.org/10.17269/cjph.104.3724>
- Daoud, N., Urquia, M. L., O'Campo, P., Heaman, M., Janssen, P. A., Smylie, J., & Thiessen, K. (2012). Prevalence of abuse and violence before, during, and after pregnancy in a national sample of Canadian women. *American Journal of Public Health, 102*, 1893–1901. <https://doi.org/10.2105/AJPH.2012.300843>
- Davies, J. M., Lyon, E., & Monti-Catania, D. (1998). *Safety planning with battered women: Complex lives/difficult choices*. Thousand Oaks, CA: Sage Publications.
- Dawes, R. The robust beauty of improper linear models in decision making. *American Psychologist, 34*(7), 571-582.
- Dawson, M., & Hotton, T. (2014). Police charging practices for incidents of intimate partner violence in Canada. *Journal of Research in Crime & Delinquency, 51*(5), 655-683. <https://doi.org/10.1177/0022427814523787>
- Day, T. (1995). *The health-related costs of violence against women in Canada: The tip of the iceberg*. London, ON: Centre for Research on Violence Against Women and Children. <https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=163639>

- Department of Justice. (2002). Spousal abuse policies and legislation. *Government of Canada*. http://www.justice.gc.ca/eng/pi/fv-vf/rep-rap/spo_e-con_a.pdf
- Department of Justice. (2015). Final report of the ad hoc federal-provincial-territorial working group reviewing spousal abuse policies and legislation. *Government of Canada*. <https://www.justice.gc.ca/eng/rp-pr/cj-jp/fv-vf/pol/p4.html>
- Department of Justice. (2016). Programming responses for intimate partner violence. *Government of Canada*. http://publications.gc.ca/collections/collection_2017/jus/J2-444-2017-eng.pdf
- Department of Justice. (2017a). An estimation of the economic impact of spousal violence in Canada, 2009. *Government of Canada*. http://justice.gc.ca/eng/rp-pr/cj-jp/fv-vf/rr12_7/p1.html
- Department of Justice. (2017b). Indigenous overrepresentation in the criminal justice system. *Justfacts*. <https://www.justice.gc.ca/eng/rp-pr/jr/jf-pf/2017/jan02.html>
- Department of Justice. (2018a). Intimate partner violence. *Government of Canada*. <https://www.justice.gc.ca/eng/rp-pr/jr/jf-pf/2017/may01.html>
- Department of Justice. (2018b). Inventory of spousal violence risk assessment tools used in Canada. *Government of Canada*. https://www.justice.gc.ca/eng/rp-pr/cj-jp/fv-vf/rr09_7/p3.html
- Department of Justice. (2021). Family violence laws. *Government of Canada*. <https://www.justice.gc.ca/eng/cj-jp/fv-vf/laws-lois.html>
- DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated Receiver Operating Characteristic curves: A

nonparametric approach. *Biometrics*, 44(3), 837-845.

<http://doi.org/10.2307/2531595>

Desmarais, S. L., Nicholls, T. L., Wilson, C. M., & Brink, J. (2012). Using dynamic risk and protective factors to predict inpatient aggression: Reliability and validity of START assessments. *Psychological Assessment*, 24(3), 685-700.

<https://doi.org/10.1037/a0026668>

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.).

Thousand Oaks, CA: Sage Publications.

deVries Robbé, M. V., Vogel, V., & Douglas, K. S. (2013). Risk factors and protective factors: A two-sided dynamic approach to violence risk assessment. *Journal of Forensic Psychiatry & Psychology*, 24(4), 440-457.

<https://doi.org/10.1080/14789949.2013.818162>

Dobash, R. E., Dobash, R. P., Cavanagh, K., & Lewis, R. (2004). Not an ordinary killer-just an ordinary guy: When men murder an intimate woman partner. *Violence Against Women*, 10(6), 577-605. <https://doi.org/10.1177%2F1077801204265015>

Dolan, M., Doyle, M., & McGovern, J. (2002) The validity of North American risk assessment tools in predicting in-patient violent behavior in England. *Legal and Criminological Psychology*, 7(2), 141-154.

<https://psycnet.apa.org/doi/10.1348/135532502760274756>

Douglas, K. S., & Skeem, J. L. (2005). Violence risk assessment: Getting specific about being dynamic. *Psychology, Public Policy, and Law*, 11(3), 347-383.

<https://psycnet.apa.org/doi/10.1037/1076-8971.11.3.347>

- Du Mont, J., Forte, T., Cohen, M. M., Hyman, I., & Romans, S. (2005). Changing help-seeking rates for intimate partner violence in Canada. *Women & Health, 41*(1), 1-19. https://doi.org/10.1300/J013v41n01_01
- Dugan, L., Nagin, D. S., & Rosenfeld, R. (2003). Do domestic violence services save lives? *National Institute of Justice Journal, 250*, 20-25. <https://www.ncjrs.gov/pdffiles1/jr000250f.pdf>
- Dunford, F. W., Huizinga, D., & Elliott, D. S. (1990). The role of arrest in domestic assault: The Omaha police experiment. *Criminology, 28*(2), 183-206. <https://doi.org/10.1111/j.1745-9125.1990.tb01323.x>
- Dunkle, K. L., Jewkes, R., Brown, H. C., Gray, G. E., McIntyre, J. A., & Harlow, S. D. (2004). Gender-based violence, relationship power, and risk of HIV in women attending antenatal clinics in South Africa. *The Lancet, 363*(9418), 1415-1421. [https://doi.org/10.1016/S0140-6736\(04\)16098-4](https://doi.org/10.1016/S0140-6736(04)16098-4)
- Dutton, D. G., & Kropp, P. R. (2000). A review of domestic violence risk instruments. *Trauma Violence and Abuse, 1*(2), 171-181. <https://doi.org/10.1177/1524838000001002004>
- Dutton, D. G., Nicholls, T. L., & Spidel, A. (2006). Female perpetrators of intimate abuse. *Canada Journal of Offender Rehabilitation, 41*(4), 1-31. doi: 10.1300/J076v41n04_01
- Dvoskin, J. A., Skeem, J. L., Novaco, R. W., & Douglas, K. S. (2012). What if psychology redesigned the criminal justice system? In J. A. Dvoskin, J. L. Skeem, R. W. Novaco, & K. S. Douglas (Eds.), *Using social science to reduce violent offending* (pp. 291–302). New York, NY: Oxford University Press.

- Eastman, N., Adshead, G., Fox, S., Latham, R., & Whyte, S. (2019). *Forensic psychiatry*. New York, NY: Oxford University Press.
- Elisha, E., Idisis, Y., Timor, U., & Addad, M. (2010). Typology of intimate partner homicide: Personal, interpersonal, and environmental characteristics of men who murdered their female intimate partner. *International Journal of Offender Therapy and Comparative Criminology*, *54*(4), 494-516.
<https://doi.org/10.1177/0306624x09338379>
- Epstein, D. (1999). Redefining the state's response to domestic violence: Past victories and future challenges. *Georgetown Journal of Gender and the Law*, *1*, 127–143.
<https://scholarship.law.georgetown.edu/cgi/viewcontent.cgi?article=2725&context=facpub>
- Eriksson, L., & Mazerolle, P. (2015). A cycle of violence? Examining family-of-origin violence, attitudes, and IPV perpetration. *Journal of Interpersonal Violence*, *30*(6), 945-964. <https://doi.org/10.1177/0886260514539759>
- Ewert v. Canada*. (2015). FC 1093
- Ewert v. Canada*. (2018). SCC 30.
- Fals-Stewart, W. (2003). The occurrence of partner physical aggression on days of alcohol consumption. *Journal of Consulting and Clinical Psychology*, *71*(1), 41–52. <https://psycnet.apa.org/doi/10.1037/0022-006X.71.1.41>
- Fals-Stewart, W., Kashdan, T. B., O'Farrell, T. J., & Birchler, G. R. (2002). Behavioral couples therapy for drug-abusing patients: Effects on partner violence. *Journal of Substance Abuse Treatment*, *22*(2), 87–96. [https://doi.org/10.1016/S0740-5472\(01\)00218-5](https://doi.org/10.1016/S0740-5472(01)00218-5)

- Farrington, D. P. (1985). Predicting self-reported and official delinquency. In D. P. Farrington & R. Tarling (Eds.), *Prediction in criminology* (pp. 150-173). New York, NY: State University of New York Press.
- Farrington, D. P. (2002). *What has been learned from self-reports about criminal careers and the causes of offending?* <https://docplayer.net/12690130-What-has-been-learned-from-self-reports-about-criminal-careers-and-the-causes-of-offending.html>
- Fidler, N. D., & Gray, K. G. (1947). *Law and the practice of nursing*. Toronto: Ryerson.
- Field, C. A., Caetano, R., & Nelson, S. (2004). Alcohol and violence related cognitive risk factors associated with the perpetration of intimate partner violence. *Journal of Family Violence, 19*(4), 249–253.
<https://doi.org/10.1023/B:JOFV.0000032635.42145.66>
- Finch, B., Gilligan, D. G., Halpin, S. A., & Valentine, M. E. (2016). The short- to medium-term predictive validity of static and dynamic risk-of-violence measures in medium- to low-secure forensic and civil inpatients. *Psychiatry, Psychology and Law, 24*(3), 410-427. <https://doi.org/10.1080/13218719.2016.1247640>
- Fleiss, J. L. (1986). The design and analysis of clinical experiments. *Biometrical Journal, 30*(3), 303-304. <https://doi.org/10.1002/bimj.4710300308>
- Flynn, S., Gask, L., Appleby, L., & Shaw, J. (2016). Homicide-suicide and the role of mental disorder: A national consecutive case series. *Social Psychiatry and Psychiatric Epidemiology, 51*(6), 877-884. <https://doi.org/10.1007/s00127-016-1209-4>

- Frank, J., Smith, B., & Novak, K. (2005). Exploring the basis of citizens' attitudes toward the police. *Police Quarterly*, 8(2), 206–228. <https://doi-org.proxy.hil.unb.ca/10.1177/1098611103258955>
- Frerichs, R., & Tuokko, H. (2006). Reliable change scores and their relation to perceived change in memory: Implications for the diagnosis of mild cognitive impairment. *Archives of Clinical Neuropsychology*, 21(1), 109-115. <https://doi.org/10.1016/j.acn.2005.08.003>
- Gabbidon, S., & Higgins, G. (2009). The role of race/ethnicity and race relations on public opinion related to the treatment of blacks by the police. *Police Quarterly*, 12(1), 102–115. <https://doi-org.proxy.hil.unb.ca/10.1177/1098611108329692>
- Garcia, L., Soria, C., & Hurwitz, E. L. (2007). Homicides and intimate partner violence: A literature review. *Trauma, Violence, & Abuse*, 8(4), 370-383. <https://doi.org/10.1177/1524838007307294>
- Garner, J. H. & Maxwell, C. D. (1999). Measuring the amount of force used by and against the police in six jurisdictions. *National Institute of Justice*. <https://www.ojp.gov/sites/g/files/xyckuh241/files/archives/ncjrs/176330-2.pdf>
- Garner, J. H., & Maxwell, C. D. (2009). Prosecution and conviction rates for intimate partner violence. *Criminal Justice Review*, 34(1), 44-79. <https://doi.org/10.1177/0734016808324231>
- Gartin, P. R. (1995). Examining differential officer effects in the Minneapolis domestic violence experiment. *American Journal of Police*, 14, 93-110. <https://doi.org/10.1108/07358549510111965>

- Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works. *Criminology*, 34(4), 575–608.
<https://doi.org/10.1111/j.1745-9125.1996.tb01220.x>
- Gerth, J., Rossegger, A., Bauch, E., & Endrass, J. (2017). Assessing the discrimination and calibration of the Ontario Domestic Assault Risk Assessment in Switzerland. *Partner Abuse*, 7(2), 168-189. <https://doi.org/10.1891/1946-6560.8.2.168>
- Gill, C. & Fitch, L. (2016). Developing and delivering a national framework for collaborative police action to intimate partner violence in Canada. *Journal of Community Safety & Well-Being*, 1(3), 1-9.
https://cacp.ca/index.html?asst_id=1200
- Gondolf, E. W. (2002). *Batterer Intervention Systems: Issues, Outcomes, and Recommendations (Sage Series on Violence Against Women)*. Thousand Oaks, CA: Sage Publications.
- Goodman-Delahunty, J. & O'Brien, K. (2014). *Reoffence risk in intrafamilial child sexual offenders*. Report to the Criminology Research Advisory Council Grant: CRG 44/10-11. Canberra: Australian Institute of Criminology.
- Goodman, L. A., Dutton, M. A., & Bennett, L. (2000). Predicting repeat abuse among arrested batterers: Use of the danger assessment scale in the criminal justice system. *Journal of Interpersonal Violence*, 15(1), 63-74.
<https://doi.org/10.1177/088626000015001005>
- Government of Canada. (2006). *Facts on alcohol and violence: Intimate partner violence and alcohol*. <https://www.canada.ca/en/public-health/services/health->

promotion/stop-family-violence/prevention-resource-centre/women/who-facts-on-alcohol-violence-intimate-partner-violence-alcohol.html

Government of Northwest Territories, Government of Nunavut, and the Legacy of Hope Foundation. (2013). *The residential school system in Canada: Understanding the past – seeking reconciliation – building hope for tomorrow*, second edition.

https://www.ece.gov.nt.ca/sites/ece/files/resources/northern_studies_10_teaching_guide.pdf

Grann, M., & Wedin, I. (2002). Risk factors for recidivism among spousal assault and spousal homicide offenders. *Psychology, Crime & Law*, 8(1), 5-23.

<https://doi.org/10.1080/10683160208401806>

Grant, S. & Rowe, M. (2011). Running the risk: Police officer discretion and family violence in New Zealand, *Policing and Society*, 21(1), 49-66.

<https://doi.org/10.1080/10439463.2010.540662>

Greaves, L., Hankivsky, O., & Kingston-Riechers, J. (1995). *Selected estimates of the costs of violence against women*. London, ON: Centre for Research on Violence Against Women and Children. http://www.learningtoendabuse.ca/our-work/pdfs/pub_greaves19951.pdf

Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical–statistical controversy. *Psychology, Public Policy, and Law*, 2(2), 293-

323. <https://doi.org/10.1037/1076-8971.2.2.293>

- Guadalupe-Diaz, X. L. (2016). Disclosure of same-sex intimate partner violence to police among lesbians, gays, and bisexuals. *Social Currents*, 3(2), 160–171. <https://doi.org/10.1177/2329496515604635>
- Guadalupe-Diaz, X. L., & Yglesias, J. (2013). Who's protected? Exploring perceptions of domestic violence law by lesbians, gays, and bisexuals. *Journal of Gay and Lesbian Social Services*, 25(4), 465–485. <https://doi.org/10.1080/10538720.2013.806881>
- Guo, B., & Harstall, C. (2008). *Spousal violence against women: Preventing recurrence*. Edmonton, AB: Institute of Health Economics.
- Hajian-Tilaki, K. (2013). Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian Journal of Internal Medicine*, 4(5), 627-635. <https://doi.org/10.1161/CIRCULATIONAHA.105.594929>
- Hannah, S. G., Turf, E. E., & Fierro, M. F. (1998). Murder-suicide in central Virginia: A descriptive epidemiologic study and empiric validation of the Hanzlick-Koponen typology. *American Journal of Forensic Medicine and Pathology*, 19(3), 275-283. <https://doi.org/10.1097/00000433-199809000-00016>
- Hanson, R. K., & Bourgon, G. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment*, 21(1), 1-21. <https://doi.org/10.1037/a0014421>
- Hanson, R. K., Bourgon, G., Helmus, L., & Hodgson, S. (2009). *A meta-analysis of the effectiveness of treatment for sexual offenders: Risk, need, and responsivity*. Ottawa: Public Safety Canada.

- Hanson, R. K., Cadsky, O., Harris, A., & Lalonde, C. (1997). Correlates of battering among 997 men: Family history, adjustment, and attitudinal differences. *Violence and Victims, 12*, 191-208. 10.1891/0886-6708.12.3.191
- Hanson, R. K., & Harris, A. J. R. (2000). Where should we intervene? Dynamic predictors of sexual offense recidivism. *Criminal Justice and Behavior, 27*(1), 6–35. <https://doi.org/10.1177/0093854800027001002>
- Hanson, R. K., & Harris, A. J. R. (2001). A structured approach to evaluating change among sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 13*(2), 105–122. <https://doi.org/10.1177/107906320101300204>
- Hanson, R. K., Harris, A. J., Scott, T. L., & Helmus, L. (2007). *Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision Project*. Ottawa, ON: Public Safety Canada.
- Hanson, R. K., Helmus, L., & Bourgon, G. (2007). *The validity of risk assessments for intimate partner violence: A meta-analysis*. Ottawa, ON: Public Safety Canada. <https://www.publicsafety.gc.ca/cnt/rsrscs/pblctns/ntmt-prtnr-vlnce/ntmt-prtnr-vlnce-eng.pdf>
- Hanson, R. K., & Morton-Bourgon, K. E. (2007). *The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis* (Corrections User Report No 2007-01). Ottawa: Public Safety and Emergency Preparedness Canada.
- Hanzlick, R., & Koponen, M. (1994). Murder-suicide in Fulton County, Georgia, 1988–1991: Comparison with a recent report and proposed typology. *The American Journal of Forensic Medicine and Pathology, 15*(2), 168-173. <https://doi.org/10.1097/00000433-199406000-00015>

- Hamberger, L. K., Lohr, J. M., Bonge, D., & Tolin, D. F. (1996). A large sample empirical typology of male spouse abusers and its relationship to dimensions of abuse. *Violence and Victims, 11*(4), 277-301. <https://doi.org/10.1891/0886-6708.11.4.277>
- Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behaviour, 20*(4), 315-335. <https://doi.org/10.1177/0093854893020004001>
- Hart, B. (1990). *Assessing whether batterers will kill*. Harrisburg, PA: Coalition Against Domestic Violence.
- Hart, S. D. (2010). The critical roles of perpetrator risk, victim vulnerability, & community support factors. Paper presented at the Second Annual Canadian Domestic Homicide Prevention Conference, Calgary, Alberta.
- Hartman, J., & Belknap, J. (2003). Beyond the gatekeepers: Court professionals' self-reported attitudes about and experiences with misdemeanor domestic violence cases. *Criminal Justice and Behavior, 30*(3), 349–373. <https://doi.org/10.1177/0093854803030003005>
- Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: The Guilford Press.
- Heckert, D. A., & Gondolf, E. W. (2004). Battered women's perceptions of risk versus risk factors and instruments in predicting repeat reassault. *Journal of Interpersonal Violence, 19*(7), 778-800. <https://doi.org/10.1177/0886260504265619>

- Heffernan, R., Wegerhoff, D. & Ward, T. (2019). Dynamic risk factors: Conceptualization, measurement, and evidence. *Aggression and Violent Behavior, 48*(1), 6-16. <https://doi.org/10.1016/j.avb.2019.06.004>
- Hegel, J., Pelletier, K. D., & Olver, M. E. (2021). Predictive properties of the Ontario Domestic Assault Risk Assessment (ODARA) in a northern Canadian prairie sample. *Criminal Justice and Behavior, 1*, 1-21. <https://doi.org/10.1177/00938548211033631>
- Heilbrun, K. (1997). Prediction versus management models relevant to risk assessment: The importance of legal decision-making context. *Law and Human Behavior, 21*(4), 347-359. <https://doi.org/10.1023/a:1024851017947>
- Helmus, L. M. (2018). Sex offender risk assessment: Where are we and where are we going? *Current Psychiatry Reports, 20*(6), 46-66. <https://doi.org/10.1007/s11920-018-0909-8>.
- Helmus, L., & Bourgon, G. (2011). Taking stock of 15 years of research on the Spousal Assault Risk Assessment Guide (SARA): A critical review. *International Journal of Forensic Mental Health, 10*(1), 64-75. <https://doi.org/10.1080/14999013.2010.551709>
- Henning, K., Jones, A., & Holdford, R. (2005). "I didn't do it, but if I did I had a good reason": Minimization, denial, and attributions of blame among male and female domestic violence offenders. *Journal of Family Violence, 20*(3), 131-139. <https://www.researchgate.net/deref/http%3A%2F%2Fdx.doi.org%2F10.1007%2Fs10896-005-3647-8>

- Henson R. K. (2001). Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. *Measurement and Evaluation in Counseling and Development, 34*(3), 177–89.
- Herbert, A., Heron, J., Barter, C., Szilassy, E., Barnes, M., Howe, L. D., Feder, G., & Fraser, A. (2021). Risk factors for intimate partner violence and abuse among adolescents and young adults: Findings from a UK population-based cohort. *Wellcome Open Research, 5*(1), 176–176.
<https://doi.org/10.12688/wellcomeopenres.16106.1>
- Hilton, Z. N. (1993). Introduction. In Z. N. Hilton (Ed.), *Legal responses to wife assault: current trends and evaluations* (pp. 3-8). Newbury Park, CA: Sage Publications.
- Hilton, Z. N., Harris, G. T., Popham, S., & Lang, C. (2010). Risk assessment among incarcerated male domestic violence offenders. *Criminal Justice and Behavior, 37*(8), 815–832. <http://dx.doi.org/10.1177/0093854810368937>
- Hilton, Z. N., Harris, G. T., & Rice, M. E. (2001). Predicting violent recidivism by serious wife assaulters. *Journal of Interpersonal Violence, 16*(5), 408-423.
<https://doi.org/10.1177/088626001016005002>
- Hilton, Z. N., Harris, G. T., & Rice, M. (2010). *Risk assessment for domestically violent men: Tools for criminal justice, offender intervention, and victim services*. Washington, D.C: American Psychological Association.
- Hilton, Z. N., Harris, G. T., Rice, M. E., Eke, A. W., & Lowe-Wetmore, T. (2007). Training front-line users in the Ontario Domestic Assault Risk Assessment (ODARA): A tool for police domestic investigations. *Canadian Journal of Police & Security Services, 5*, 92-96. <https://www.researchgate.net/profile/Grant>

_Harris/publication/230601461_Training_front-line_users_in_the_Ontario_Domestic_Assault_Risk_Assessment_ODARA_a_tool_for_police_domestic_investigations/links/00b495238a01e2cff4000000/Training-front-line-users-in-the-Ontario-Domestic-Assault-Risk-Assessment-ODARA-a-tool-for-police-domestic-investigations.pdf

Hilton, Z. N., Harris, G. T., Rice, M. E., Lang, C., Cormier, C. A., & Lines, K. J. (2004).

A brief actuarial assessment for the prediction of wife assault recidivism: The Ontario Domestic Assault Risk Assessment. *Psychological Assessment, 16*(3), 267-275. <https://doi.org/10.1037/1040-3590.16.3.267>

Hilton, Z. N., Popham, S., Lang, C., & Harris, G. T. (2014). Preliminary validation of the ODARA for female intimate partner violence offenders. *Partner Abuse, 5*(2), 189–203. <http://dx.doi.org/10.1891/1946-6560.5.2.189>

Hirschel, J. D., Hutchison, I. W., & Dean, C. W. (1992). The failure of arrest to deter spouse abuse. *Journal of Research in Crime and Delinquency, 29*(1), 7-33. <https://doi.org/10.1177/0022427892029001002>

Ho, T. N. (2000). Domestic violence in a southern city: The effects of a mandatory arrest policy on male-versus-female aggravated assault incidents. *American Journal of Criminal Justice, 25*, 107-118. <https://doi.org/10.1007/BF02886814>

Hoge, R. D., & Andrews, D. A. (1996). *Assessing the youthful offender: Issues and techniques*. New York: Plenum.

Holtzworth-Munroe, A., & Hutchinson, G. (1993). Attributing negative intent to wife behavior: The attributions of maritally violent versus nonviolent men. *Journal of*

Abnormal Psychology, 102(2), 206–211. <https://doi.org/10.1037//0021-843x.102.2.206>

Holtzworth-Munroe, A., Meehan, J. C., Herron, K., Rehman, U., & Stuart, G. L. (2000). Testing the Holtzworth-Munroe and Stuart typology. *Journal of Consulting and Clinical Psychology*, 68(6), 1000-1019. <https://doi.org/10.1037//0022-006x.68.6.1000>

Holtzworth-Munroe, A., Meehan, J. C., Herron, K., Rehman, U., & Stuart, G. L. (2003). Do subtypes of maritally violent men continue to differ over time? *Journal of Consulting and Clinical Psychology*, 71(4), 728-740. <https://doi.org/10.1037/0022-006x.71.4.728>

Holtzworth-Munroe, A. & Stuart, G. L. (1994). Typologies of male batterers: Three subtypes and the differences among them. *Psychological Bulletin*, 116(3), 476-497. <https://doi.org/10.1037/0033-2909.116.3.476>

Horwitz, S. H., Mitchell, D., Larussa-Trott, M., Santiago, L., Pearson, J., Skiff, D. M., & Cerulli, C. (2011). An inside view of police officers' experience with domestic violence. *Journal of Family Violence*, 26(8), 617–625. <https://doi.org/10.1007/s10896-011-9396-y>

Hudson, S. M., Wales, D. S., Bakker, L., & Ward, T. (2002). Dynamic risk factors: The Kia Marama evaluation. *Sexual Abuse: A Journal of Research and Treatment: Official Journal of the Association for the Treatment of Sexual Abusers*, 14(2), 103-119. <https://doi.org/10.1177/107906320201400203>

Humphrey, C., Thiara, R. K., Regan, L., Lovet, J., Kennedy, L., & Gibson, A. (2005). *Prevention not Prediction? A Preliminary Evaluation of the Metropolitan Police*

Domestic Violence Risk Assessment Model. London: Centre for the Study of Safety and Wellbeing, University of Warwick and Child and Woman Abuse Studies Unit. https://cwasu.org/wp-content/uploads/2016/07/ACPO_PreventionnotPrediction2005.pdf

Huss, M. T., & Langhinrichsen-Rohling, J. (2006). Assessing the generalization of psychopathy in a clinical sample of domestic violence perpetrators. *Law and Human Behavior, 30*(5), 571-586. <https://doi.org/10.1007/s10979-006-9052-x>

Huss, M. T., & Ralston, A. (2008). Do batterer subtypes actually matter? Treatment completion, treatment response, and recidivism across a batterer typology. *Criminal Justice and Behavior, 35*(6), 710-724. <https://www.researchgate.net/derf/http%3A%2F%2Fdx.doi.org%2F10.1177%2F0093854808316218>

Iacobucci, D. (2012). Mediation analysis and categorical variables: The final frontier. *Journal of Consumer Psychology, 22*(4), 582-594. <https://doi.org/10.1016/j.jcps.2012.03.006>

Ibrahim, D. (2020). Public perceptions of the police in Canada's provinces, 2019. *Juristat, Statistics Canada*. (Catalogue no. 85-002-X). <https://www150.statcan.gc.ca/n1/pub/85-002-x/2020001/article/00014-eng.htm>

International Association of Chiefs of Police. (2015). *Law enforcement's role in victim compensation*. <https://www.theiacp.org/LE-Role-in-Victim-Compensation>

Ishoy, G. (2016). The theory of planned behavior and policing: How attitudes about behavior, subjective norms, and perceived behavioral control affect the discretionary enforcement decisions of police officers. *Criminal Justice Studies, 29*(1), 1-18. doi: 10.1080/1478601X.2016.1225362

- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12–19. <https://doi.org/10.1037//0022-006x.59.1.12>
- Jaffe, P., Hastings, E., Reitzel, D., & Austin, G. (1993). The impact of police laying charges. In Z. N. Hilton (Ed.), *Legal Responses to Wife Assault: Current Trends and Evaluations* (pp. 62-95). Newbury Park, CA: Sage Publications.
- Jaffe, P., Wolfe, D. A., Telford, A., & Austin, G. (1986). The impact of police charges in incidents of wife abuse. *Journal of Family Violence, 1*, 37-49. <https://doi.org/10.1007/BF00977031>
- Johnson, H., & Dawson, M. (2011). *Violence against women in Canada: Research and policy perspectives*. Don Mills, ON: Oxford University Press Canada.
- Jung, S., & Buro, K. (2017). Appraising risk for intimate partner violence in a police context. *Criminal Justice and Behavior, 44*(2), 240-260. <https://psycnet.apa.org/doi/10.1177/0093854816667974>
- Kalyal, H. (2021). One person's evidence is another person's nonsense: Why police organizations resist evidence-based practices. *Policing: A Journal of Policy and Practice, 14*(4), 1151–1165. <https://doi.org/10.1093/police/pay106>
- Kaplan, E. L. and Meier, P. (1958) Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association, 53*, 457-481. <http://dx.doi.org/10.1080/01621459.1958.10501452>
- Kaukinen, C. (2004). The help-seeking strategies of female violent-crime victims: The direct and conditional effects of race and the victim-offender

- relationship. *Journal of Interpersonal Violence*, 19(9), 967–990.
<https://doi.org/10.1177/0886260504268000>
- Kebbell, M., Hurren, E., Mazerolle, P., & Australian Institute of Criminology. (2006). *An investigation into the effective and ethical interviewing of suspected sex offenders*. Canberra: Australian Institute of Criminology.
- Kendall-Tackett, K. A., & Giacomoni, S. M. (2007). *Intimate partner violence*. Kingston, NJ: Civic Research Institute.
- Kerr, R., & McLean, J. (1996). *Paying for violence: Some of the costs of violence against women in B.C.* Victoria, BC: The Ministry.
http://www.cwhn.ca/en/www.llbc.leg.bc.ca/public/PubDocs/bcdocs/257780/paying_for_violence.pdf
- Klein, A. R., & Tobin, T. (2008). A longitudinal study of arrested batterers, 1995-2005: Career criminals. *Violence Against Women*, 14(2), 136-157.
<https://doi.org/10.1177/1077801207312396>
- Klepfishz, G., Daffern, M., & Day, A. (2016). Understanding dynamic risk factors for violence. *Psychology, Crime, & Law*, 22(2), 124-137.
<https://doi.org/10.1080/1068316X.2015.1109091>
- Klostermann, K. C. (2006). Substance abuse and intimate partner violence: Treatment considerations. *Substance Abuse Treatment, Prevention, and Policy*, 1(1), 24–32.
<https://doi.org/10.1186/1747-597X-1-24>
- Koehler, J. A., Lösel Friedrich, Akoensi, T. D., & Humphreys, D. K. (2013). A systematic review and meta-analysis on the effects of young offender treatment

- programs in Europe. *Journal of Experimental Criminology*, 9(1), 19–43.
<https://doi.org/10.1007/s11292-012-9159-7>
- Koss, M. P. (1990). The women's mental health research agenda: Violence against women. *American Psychologist*, 45(3), 374-380. <https://doi.org/10.1037//0003-066x.45.3.374>
- Koziarski, J., & Kalyal, H. (2020). Resistance to evidence-based policing: Canadian police executives' perceptions as to which level of Canadian policing is most resistant. *Police Practice and Research*, 22(1), 1–14.
<https://doi.org/10.1080/15614263.2020.1786690>
- Krishnan, S. P., Hilbert, J. C., & VanLeeuwen, D. (2001). Domestic violence and help-seeking behaviors among rural women: Results from a shelter-based study. *Family & Community Health*, 24(1), 28–38. doi: 10.1097/00003727-200104000-00006
- Kropp P. R. (2008). Intimate partner violence risk assessment and management. *Violence and victims*, 23(2), 202–220. <https://doi.org/10.1891/0886-6708.23.2.202>
- Kropp, P. R., & Hart, S. D. (2000). The Spousal Assault Risk Assessment (SARA) Guide: Reliability and validity in adult male offenders. *Law and Human Behavior*, 24(1), 101-118. <https://doi.org/10.1023/a:1005430904495>
- Kropp, P. R., & Hart, S. D. (2004). *The development of the Brief Spousal Assault Form for the Evaluation of Risk (B-SAFER): A tool for criminal justice professionals*. Ottawa, ON: Department of Justice (Government of Canada).

- Kropp, P. R., & Hart, S. D. (2015). *The Spousal Assault Risk Assessment Guide Version 3 (SARA-V3)*. Vancouver, BC: ProActive ReSolutions.
- Kropp, P. R., Hart, S. D., Webster, C. D., & Eaves, D. (1995). *Manual for the Spousal Assault Risk Assessment Guide* (2nd ed.). Vancouver, BC: Institute on Family Violence.
- Kropp, P. R., Hart, S. D., Webster, C. D., & Eaves, D. (1999). *Manual for the Spousal Assault Risk Assessment Guide* (3rd ed.). Toronto, ON: Multi-Health Systems.
- Krug, E. G., Dahlberg, L. L., Mercy, J. A., Zwi, A. B., & Lozano, R. (2002). *World Report on Violence and Health*. Geneva: World Health Organization.
http://whqlibdoc.who.int/publications/2002/9241545615_eng.pdf
- Kubik, W., Bourassa, C., & Hampton, M. (2009). Stolen sisters, second class citizens, poor health: The legacy of colonization in Canada. *Humanity & Society*, 33, 18–34. <https://doi.org/10.1177/016059760903300103>
- Labrecque, R. M., Smith, P., Lovins, B. K., & Latessa, E. J. (2014). The importance of reassessment: How changes in the LSI-R risk score can improve the prediction of recidivism. *Journal of Offender Rehabilitation*, 53(2), 116-128.
<https://doi.org/10.1080/10509674.2013.868389>
- Lai, Y. L., & Zhao, J. S. (2010). The impact of race/ethnicity, neighborhood context, and police/citizen interaction on residents' attitudes toward the police. *Journal of Criminal Justice*, 38(4), 685–692. <https://doi.org/10.1016/j.jcrimjus.2010.04.042>
- Lamis, D., Leenaars, L., Jahn, D., & Lester, D. (2013). Intimate partner violence: Are perpetrators also victims and are they more likely to experience suicide

ideation? *Journal of Interpersonal Violence*, 28(16), 3109-3109.

<https://doi.org/10.1177/0886260513488691>

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-74. <https://www.jstor.org/stable/2529310>

Langhinrichsen-Rohling, J., Huss, M. T., & Ramsey, S. (2000). The clinical utility of batterer typologies. *Journal of Family Violence*, 15, 37-53.

<https://doi.org/10.1023/A:1007597319826>

LaPrairie, C. (1995). Seen but not heard: Native people in four Canadian inner cities. *The Journal of Human Justice*, 6, 30-45.

<https://doi.org/10.1007/BF02585441>

Lauria, I., McEwan, T. E., Luebbers, S., Simmons, M., & Ogloff, J. R. P. (2017). Evaluating the Ontario Domestic Assault Risk Assessment in an Australian frontline police setting. *Criminal Justice and Behavior*, 44(12), 1545- 1558.

<https://doi.org/10.1177/0093854817738280>

Lawson, D. M., Weber, D., Beckner, H. M., Robinson, L., Marsh, N., & Cool, A. (2003). Men who use violence: Intimate violence versus non-intimate violence. *Violence and Victims*, 18(3), 259-277. <https://doi.org/10.1891/vivi.2003.18.3.259>

Legislative Assembly of New Brunswick. (2018). *Intimate partner violence intervention act*. Government of New Brunswick.

<https://www.gnb.ca/legis/bill/FILE/58/3/Bill-47-e.htm>

Leisenring, A. (2008). Controversies surrounding mandatory arrest policies and the police response to intimate partner violence. *Sociology Compass*, 2(2), 451-466.

<https://doi.org/10.1111/j.1751-9020.2008.00095.x>

- Leonard, K. E., & Senchak, M. (1996). Prospective prediction of husband marital aggression within newlywed couples. *Journal of Abnormal Psychology, 105*(3), 369–380. <https://doi.org/10.1037//0021-843x.105.3.369>
- Lewis, K., Olver, M. E., & Wong, S. P. (2013). The Violence Risk Scale: Predictive validity and linking changes in risk with violent recidivism in a sample of high-risk offenders with psychopathic traits. *Assessment, 20*(2), 150-164. <https://doi.org/10.1177/1073191112441242>
- Lipsey, M. W. (2009). The primary factors that characterize effective interventions with juvenile offenders: a meta-analytic overview. *Victims & Offenders, 4*(2), 124–147. <https://doi.org/10.1080/15564880802612573>
- Litwack, T. R., & Schlesinger, L. B. (1999). Dangerousness risk assessments: Research, legal and clinical considerations. In A. K. Hess & I. B. Weiner (Eds.), *Handbook of Forensic Psychology* (2nd ed., pp. 171-217). New York, NY: Wiley.
- Litwin, M. S. (2003). *How to assess and interpret survey psychometrics* (2nd ed). Thousand Oaks, CA: Sage Publications.
- Loinaz, I. (2014). Typologies, risk and recidivism in partner-violent men with the B-SAFER: A pilot study. *Psychology, Crime & Law, 20*(2), 183-198. <http://dx.doi.org/10.1080/1068316X.2013.770854>
- Lowenkamp, C. T., Latessa, E. J., & Holsinger, A. M. (2006). The risk principle in action: What have we learned from 13,676 offenders and 97 correctional programs? *Crime & Delinquency, 52*(1), 77–93. <https://doi.org/10.1177/0011128705281747>

- Lum, C., Telep, C. W., Koper, C. S., & Grieco, J. (2012). Receptivity to research in policing. *Justice Research and Policy, 14*(1), 61–95.
<https://doi.org/10.3818/JRP.14.1.2012.61>
- Luke, D. A., & Homan, S. M. (1998). Time and change: Using survival analysis in clinical assessment and treatment evaluation. *Psychological Assessment, 10*(4), 360-378. <https://psycnet.apa.org/doi/10.1037/1040-3590.10.4.360>
- Lysova, A., Dim, E., & Dutton, D. (2019). Prevalence and consequence of intimate partner violence in Canada as measured by the national victimization survey. *Partner Abuse, 10*(2), 199-221. <https://doi:10.1891/1946-6560.10.2.199>
- Lysova, A., Hanson, K., Dixon, L., Douglas, E. M., Hines, D. A., & Celi, E. M. (2020). Internal and external barriers to help seeking: Voices of men who experienced abuse in the intimate relationships. *International Journal of Offender Therapy and Comparative Criminology, 5*, 1-22.
<https://doi.org/10.1177/0306624X20919710>
- MacKinnon, D. P., & Dwyer, J. H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review, 17*(2), 144–158.
<https://doi.org/10.1177%2F0193841X9301700202>
- MacKinnon, D., Lockwood, C., Brown, C., Wang, W., & Hoffman, J. (2007). The intermediate endpoint effect in logistic and probit regression. *Clinical Trials, 4*(5), 499-513. <https://dx.doi.org/10.1177%2F1740774507083434>
- Martin, D. I. (1998). Retribution revisited: A reconsideration of feminist criminal law reform strategies. *Osgoode Hall Law Journal, 36*(1), 151-188.

<http://digitalcommons.osgoode.yorku.ca/cgi/viewcontent.cgi?article=1570&context=ohlj>

Maynard, R. (2017). *Policing black lives: State violence in Canada from slavery to the present*. Fernwood Publishing.

McCuaig, G. (2012). *British Columbia Charge Assessment Review*. Vancouver, BC: BC Justice Reform Initiative. <https://www2.gov.bc.ca/assets/gov/law-crime-and-justice/criminal-justice/prosecution-service/reports-publications/2012-bc-charge-assessment-review-mccuaig-report.pdf>

McDougall, A. M., (2014). *Understanding factors that impact responsivity within case management plans of community-based offenders with and without mental health needs*. Unpublished doctoral dissertation. Saint John, NB: University of New Brunswick.

McFarlane, J., Soeken, K., Reel, S., Parker, B., & Silva, C. (1997). Resource use by abused women following an intervention program: Associated severity of abuse and reports of abuse ending. *Public Health Nursing, 14*(4), 244–50. <https://doi-org.proxy.hil.unb.ca/10.1111/j.1525-1446.1997.tb00297.x>

McLaughlin, E. M., & Rozee, P. D. (2001). Knowledge about heterosexual versus lesbian battering among lesbians. *Women & Therapy, 23*(3), 39–58. doi: 10.1300/J015v23n03_04

McMillan, J. H. & Schumacher, S. (2001). *Research in education: A conceptual introduction*. New York, NY: Longman.

McNiel, D. E., Gregory, A. L., Lam, J. N., Binder, R. L., & Sullivan, G. R. (2003). Utility of decision support tools for assessing acute risk of violence. *Journal of*

Consulting and Clinical Psychology, 71(5), 945-953.

<https://doi.org/10.1037/0022-006x.71.5.945>

Mechanic, M. B., Weaver, T. L., & Resick, P. A. (1996). Mental health consequences of intimate partner abuse: A multidimensional assessment of four different forms of abuse. *Violence Against Women*, 14(6), 634–654.

<https://doi.org/10.1177/1077801208319283>

Medeiros, R. A., & Straus, M. A. (2006). Risk factors for physical violence between dating partners: Implications for gender inclusive prevention and treatment of family violence. In J. Hamel & T. Nicholls (Eds.), *Family Interventions in Domestic Violence: A Handbook of Gender-Inclusive Theory and Treatment* (pp. 59-86). Springer Publishing Company, New York, NY.

Ménard, K. S., Anderson, A. L., & Godboldt, S. M. (2009). Gender differences in intimate partner recidivism: A 5-year follow-up. *Criminal Justice and Behaviour*, 36(1), 61-76. doi:10.1177/0093854808325905

Mental Health Centre Penetanguishene Research Department & Ontario Provincial Police Behavioural Sciences Section (2005). *Ontario Domestic Assault Risk Assessment: General Scoring Criteria Revised Edition*. Unpublished training manual.

Messing, J. T., & Thaller, J. The average predictive validity of intimate partner violence risk assessment instruments. *Journal of Interpersonal Violence*, 28(7), 1537-1558. <https://doi.org/10.1177/0886260512468250>

Migliaccio, T. A. (2001). Marginalizing the battered male. *Journal of Men's Studies*, 9(2), 205-226. <https://doi.org/10.3149/jms.0902.205>

- Migliaccio, T. A. (2002). Abused husbands: a narrative analysis. *Journal of Family Issues*, 23(1), 26–52. <https://doi.org/10.1177/0192513X02023001002>
- Millar, A. (2009). Inventory of spousal violence risk assessment tools used in Canada. Department of Justice. <http://onlinetraining.learningtoendabuse.ca/sites/default/files/lessons/Inventory%20of%20Spousal%20Violence%20Risk%20Assessment%20Tools%20Used%20in%20Canada.pdf>
- Millar, A., Code, R., & Ha, L. (2013). *Inventory of spousal violence risk assessment tools used in Canada*. Ottawa, ON: Department of Justice. http://www.justice.gc.ca/eng/rp-pr/cj-jp/fv-vf/rr09_7/rr09_7.pdf
- Miller, K-L., & Du Mont, J. (2000). Countless abused women: Homeless and inadequately housed. *Canadian Woman Studies*, 20(3), 115-122. <https://cws.journals.yorku.ca/index.php/cws/article/viewFile/12674/11757>
- Mills, J., Kroner, D., & Hemmati, T. (2004). The measures of criminal attitudes and associates (MCAA). *Criminal Justice and Behavior*, 31(6), 717-733. doi: 10.1177/0093854808322239
- Mills, L. G. (2003). *Insult to injury: Rethinking our responses to intimate abuse*. Princeton, NJ: Princeton University Press.
- Moffitt, T. E., Caspi, A., Krueger, R. F., Magdol, L., Margolin, G., Silva, P. A., & Sydney, R. (1997). Do partners agree about abuse in their relationship?: A psychometric evaluation of interpartner agreement. *Psychological Assessment*, 9(1), 47-56. <https://doi.org/10.1037/1040-3590.9.1.47>
- Monahan, J. (1981). *Predicting violent behaviour: An assessment of clinical techniques*. Beverly Hills, CA: Sage.

- Monahan, J., Steadman, H. J., Silver, E., Appelbaum, P. S., Robbins, P., & Mulvey, E. P. (2001). *Rethinking risk assessment: The MacArthur study of mental disorder and violence*. New York, NY: Oxford University Press.
- Moore, T. M., Elkins, S. R., McNulty, J. K., Kivisto, A. J., & Handsel, V. A. (2011). Alcohol use and intimate partner violence perpetration among college students: Assessing the temporal association using electronic diary technology. *Psychology of Violence, 1*(4), 315-328. <https://doi.org/10.1037/a0025077>
- Morley, S., & Dowzer, C.N. (2014). *Manual for the Leeds Reliable Change Indicator: Simple Excel applications for the analysis of individual patient and group data*. Leeds, UK: University of Leeds.
- Moser, A. E. (2012). *Validation and modification of the Ontario Domestic Assault Risk Assessment (ODARA) instrument: An early warning system*. Unpublished master's thesis. Saint John, NB: University of New Brunswick.
- Murphy, C. M., Morrel, T. M., Elliott, J. D., & Neavins, T. M. (2003). A prognostic indicator scale for the treatment of partner abuse perpetrators. *Journal of Interpersonal Violence, 18*(9), 1087-1105. <https://doi.org/10.1177/0886260503254515>
- Murphy, C. M., Winters, J., O'Farrell, T. J., Fals-Stewart, W., & Murphy, M. (2005). Alcohol consumption and intimate partner violence by alcoholic men: Comparing violence and nonviolent conflicts. *Psychology of Addictive Behaviors, 19* (1), 35–42. <https://doi.org/10.1037/0893-164X.19.1.35>
- New Brunswick Multicultural Council (2018). *Domestic and intimate partner violence against immigrant women in New Brunswick: Provincial reference guide*.

http://www.nb-mc.ca/wp-content/uploads/2019/09/NBMC_Provincial-Reference-Guide_ENG_Web.pdf

- Nicholls, T. L., Desmarais, S. L., Douglas, K. S., & Kropp, P. R. (2006). Violence risk assessments with perpetrators of intimate partner abuse. In J. Hamel & T. Nicholls (Eds.), *Family Interventions in Domestic Violence: A Handbook of Gender-Inclusive Theory and Treatment* (pp. 275-301). New York, NY: Springer Publishing.
- Nicholls, T. L., Pritchard, M. M., Reeves, K. A., & Hilterman, E. (2013). Risk assessment in intimate partner violence: A systematic review of contemporary approaches. *Partner Abuse, 4*(1), 76–168.
<https://psycnet.apa.org/doi/10.1891/1946-6560.4.1.76>
- Nimmo, B. (2012). *Police family violence response: Strengthening systems to enhance prevention and intervention*. The Family Violence Symposium: Wellington, New Zealand. <https://nzfvc.org.nz/content/strengthening-systems-eliminate-family-violence-family-violence-symposium-2012>
- Northcott, M. (2012). Intimate partner violence risk assessment tools: A review. *Department of Justice Canada*. http://www.justice.gc.ca/eng/rp-pr/cj-jp/fv-vf/rr12_8/rr12_8.pdf
- O’Keefe, M. (1998). Factors mediating the link between witnessing interparental violence and dating violence. *Journal of Family Violence, 13*(1), 39–57.
<https://doi.org/10.1023/A:1022860700118>
- O’Leary, K. D. (1999). Psychological abuse: A variable deserving critical attention in domestic violence. *Violence and Victims, 14*, 3–23.

- Office of Attorney General. (2017). Pre-charge screening. *Government of New Brunswick*. <https://www2.gnb.ca/content/dam/gnb/Departments/ag-pg/PDF/en/PublicProsecutionOperationalManual/Policies/Pre-chargeScreening.pdf>
- Office of the Chief Coroner. (2015). Recommendations from the domestic violence death review committee. *Government of New Brunswick*. <https://www.mmiwg-ffada.ca/wp-content/uploads/2019/05/21-Domestic-Violence-Death-Committee-Dec-2015.pdf>
- Ogle, R. S., & Jacobs, S. (2002). *Self-defense and battered women who kill: A new framework*. Westport, CT: Praeger Publishers.
- Ogles, B. M., Lunnen, K. M., & Bonesteel, K. (2001). Clinical significance: History, application, and current practice. *Clinical Psychology Review, 2*(3), 421–446. [https://doi.org/10.1016/S0272-7358\(99\)00058-6](https://doi.org/10.1016/S0272-7358(99)00058-6)
- Olver, M. E., & Jung, S. (2017). Incremental prediction of intimate partner violence: An examination of three risk measures. *Law and Human Behavior, 1*(5), 1-14. <https://doi.org/10.1037/lhb0000251>
- Olver, M. E., Nicholaichuk, T. P., Kingston, D. A., & Wong, S. P. (2014). A multisite examination of sexual violence risk and therapeutic change. *Journal of Consulting and Clinical Psychology, 82*(2), 312-324. <https://doi.org/10.1037/a0035340>
- Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2014). Thirty years of research on the level of service scales: A meta-analytic examination of predictive accuracy and sources of variability. *Psychological Assessment, 26*(1), 156–176. <https://doi.org/10.1037/a0035080>

- Outlaw, M. (2009). No one type of intimate partner abuse: Exploring physical and non-physical abuse among intimate partners. *Journal of Family Violence, 24*(4), 263-272. <https://doi.org/10.1007/s10896-009-9228-5>
- Owusu-Bempah, A. (2017). Race and policing in historical context: Dehumanization and the policing of black people in the 21st century. *Theoretical Criminology, 21*(1), 23-34. <https://doi-org.proxy.hil.unb.ca/10.1177/1362480616677493>
- Palmer, I. (2011). *Is the United Kingdom Police Service receptive to evidence-based policing? Testing attitudes towards experimentation*. Unpublished Master's thesis. Cambridge, England: University of Cambridge.
- Palmer, I., Kirby, S., & Coleman, R. (2019). Assessing the appetite for evidence based policing: A UK based study. *International Journal of Police Science & Management, 21*(2), 91–100. <https://doi.org/10.1177/1461355719838930>
- Pan, H. S., Neidig, P. H., & O'Leary, K. D. (1994). Predicting mild and severe husband-to-wife physical aggression. *Journal of Consulting and Clinical Psychology, 62*(5), 975 -981. <https://doi.org/10.1037//0022-006x.62.5.975>
- Park, H. A. (2013). An introduction to logistic regression: From basic concepts to interpretation with particular attention to nursing domain. *Journal of Korean Academy of Nursing, 43*(2), 154-164. <https://doi.org/10.4040/jkan.2013.43.2.154>
- Park, S. H., Goo, J. M., & Jo, C.-H. (2004). Receiver operating characteristic (ROC) curve: Practical review for radiologists. *Korean Journal of Radiology, 5*(1), 11–18. <https://dx.doi.org/10.3348%2Fkjr.2004.5.1.11>
- Parks, K. A., Hseih, Y., Bradizza, C. M., & Romosz, A. M. (2008). Factors influencing the temporal relationship between alcohol consumption and experiences with

- aggression among college women. *Psychology of Addictive Behaviors*, 22(2), 210–218. <https://doi.org/10.1037/0893-164X.22.2.210>
- Parmar, M. K., & Machin, D. (1995). *Survival Analysis: A Practical Approach*. New York, NY: John Wiley & Sons.
- Pattavina, A., Hirschel, D., Buzawa, E., Faggiani, D., & Bentley, H. (2007). A comparison of the police response to heterosexual versus same-sex intimate partner violence. *Violence against Women*, 13(4), 374–394. <https://doi.org/10.1177/1077801207299206>
- Pearlman, D. N., Zierler, S., Gjelsvik, A., & Verhoek-Oftedahl, W. (2003). Neighborhood environment, racial position, and risk of police-reported domestic violence: A contextual analysis. *Public Health Reports*, 118(1), 44–58. [https://doi.org/10.1016/S0033-3549\(04\)50216-9](https://doi.org/10.1016/S0033-3549(04)50216-9)
- 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(12), 1373–1379. [https://doi.org/10.1016/s0895-4356\(96\)00236-3](https://doi.org/10.1016/s0895-4356(96)00236-3)
- Pernanen, K., Cousineau, M.-M., Brochu, S., & Sun, F. (2002). Proportions of crimes associated with alcohol and other drugs in Canada. *Canadian Centre on Substance Abuse*. <https://www.ccsa.ca/sites/default/files/2019-04/ccsa-009105-2002.pdf>
- Peterson, J., & Strand, S. (2017). Recidivism in intimate partner violence among antisocial and family-only perpetrators. *Criminal Justice and Behavior*, 44(11), 1477–1495. <https://doi.org/10.1177/0093854817719916>

- Pettersson, J., Strand, S., & Selenius, H. (2016). Risk factors for intimate partner violence: A comparison of antisocial and family-only perpetrators. *Journal of Interpersonal Violence, 1*(2), 1-21. <https://doi.org/10.1177/0886260516640547>
- Peterson, D. S., & Schroeder, J. A. (2019). *Domestic violence in international context*. London: Routledge.
- Piquero, A. R., Brame, R., Fagan, J., & Moffitt, T. E. (2006). Assessing the offending activity of criminal domestic violence suspects: Offense specialization, escalation, and de-escalation evidence from the Spouse Assault Replication Program. *Public Health Reports, 121*(4), 409–418. <https://www.jstor.org/stable/20056982>
- Paoline, E.A., 2003. Taking stock: Toward a richer understanding of police culture. *Journal of Criminal Justice, 31*, 199-214.
- Pratt-Eriksson, D., Bergbom, I., & Lyckhage, E. D. (2014). Don't ask don't tell: Battered women living in Sweden encounter with healthcare personnel and their experience of the care given. *International Journal of Qualitative Studies on Health and Well-Being, 9*(1), 1-7. <https://doi.org/10.3402/qhw.v9.23166>
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*, 879-891.
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers, 36*, 717-731.

Prendergast, M. L., Pearson, F. S., Podus, D., Hamilton, Z. K., & Greenwell, L. (2013).

The Andrews' principles of risk, need, and responsivity as applied in drug abuse treatment programs: Meta-analysis of crime and drug use outcomes. *Journal of Experimental Criminology*, 9(3), 275–300. <https://doi.org/10.1007/s11292-013-9178-z>

Province of New Brunswick (2012). *Moving from theory to outcomes: New Brunswick's crime prevention and reduction strategy*. <https://www2.gnb.ca/content/dam/gnb/Departments/ps-sp/pdf/Publications/MovingFromTheoryToOutcomes.pdf>

Provincial Office of Domestic Violence. (2014). British Columbia's provincial domestic violence plan. *Government of British Columbia*. https://www2.gov.bc.ca/assets/gov/public-safety-and-emergency-services/domestic-violence/podv/provincial_domestic_violence_plan.pdf

Public Safety Canada. (2015). National crime prevention centre. *Government of Canada*. <https://www.publicsafety.gc.ca/cnt/cntrng-crm/crm-prvntn/ntnl-crm-prvntn-cntr-en.aspx>

Public Safety Canada. (2017). Domestic violence death review. *Government of Canada*. <https://www2.gnb.ca/content/gnb/en/departments/public-safety/law-enforcement-and-inspections/content/coroner-services/domestic-violence-death-review-committee.html>

Quinsey, V. L., & Ambtman, R. (1979). Variables affecting psychiatrists' and teachers' assessments of the dangerousness of mentally ill offenders. *Journal of Consulting and Clinical Psychology*, 47(2), 353-62. <https://doi.org/10.1037//0022-006x.47.2.353>

- Quinsey, V. L., Harris, G. T., Rice, G. T., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.
- Quinsey, V. L., Harris, G. T., Rice, G. T., & Cormier, C. A. (2006). *Violent offenders: Appraising and managing risk* (2nd ed.). Washington, DC: American Psychological Association.
- 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*(5), 615-620.
<https://doi.org/10.1007/s10979-005-6832-7>
- Ristock, J. L. (2001). Decentering heterosexuality: Responses of feminist counselors to abuse in lesbian relationships. *Women & Therapy, 23*(3), 59–59. doi: 10.1300/J015v23n03_05
- Ristock, J. L. (2003). Exploring dynamics of abusive lesbian relationships: Preliminary analysis of a multi-site, qualitative study. *American Journal of Community Psychology, 31*(3/4), 329–341. doi: 10.1023/A:1023971006882
- Robertson, K., & Murachver, T. (2009). Attitudes and attributions associated with female and male partner violence. *Journal of Applied Social Psychology, 39*(7), 1481-1512. <https://psycnet.apa.org/doi/10.1111/j.1559-1816.2009.00492.x>
- Rodgers K. (1994). Wife assault: The findings of a national survey. *Juristat, Statistics Canada, 14*(9), 1–21.
- Roebuck, B., McGlinchey, D., Hastie, K., Taylor, M., Roebuck, M., Bhele, S., Hudson, E., & Xavier, R. G. (2020). Male survivors of intimate partner violence in

Canada. *Office of the Federal Ombudsman for Victims of Crime*.

<https://www.victimsfirst.gc.ca/res/cor/ipv-ipv/index.html>

Roehl, J., & Guertin, K. (2000). Intimate partner violence: The current use of risk assessments in sentencing offenders. *The Justice System Journal, 21*(2), 171-198.
<https://www.jstor.org/stable/27977017>

Rogers, M. J., & Follingstad, D. R. (2014). Women's exposure to psychological abuse: Does experience predict mental health outcomes? *Journal of Family Violence, 29*(6), 595–611. <https://psycnet.apa.org/doi/10.1007/s10896-014-9621-6>

Rollè L, Giardina, G., Caldarera, A. M., Gerino, E., & Brustia, P. (2018). When intimate partner violence meets same sex couples: A review of same sex intimate partner violence. *Frontiers in Psychology, 9*(1), 1506–1506.
<https://doi.org/10.3389/fpsyg.2018.01506>

Room, R., Babor, T., & Rehm, J. (2005). Alcohol and public health. *Lancet, 365*(1), 5-11.

Rosenbaum, A. & O'Leary, K. D. (1981). Marital violence: Characteristics of abusive couples. *Journal of Consulting and Clinical Psychology, 49*(1), 63–71.
<https://psycnet.apa.org/doi/10.1037/0022-006X.49.1.63>

Rotenberg, C. (2019). Police-reported violent crimes against young women and girls in Canada's Provincial North and Territories, 2017. *Juristat*.
<https://www150.statcan.gc.ca/n1/pub/85-002-x/2019001/article/00012-eng.htm>

Royal Canadian Mounted Police. (2019). Intimate partner violence and abuse. *Government of Canada*. <https://www.rcmp-grc.gc.ca/en/intimate-partner-violence-and-abuse>

- Russell, B., & Sturgeon, J. A. (D. (2019). Police evaluations of intimate partner violence in heterosexual and same-sex relationships: Do experience and training play a role? *Journal of Police and Criminal Psychology: The Official Journal of the Society for Police and Criminal Psychology*, 34(1), 34–44.
<https://doi.org/10.1007/s11896-018-9279-8>
- Sapers, H. (2015). *Annual report of the office of the correctional investigator 2014-2015*. Office of the Correctional Investigator. <https://www.ocibec.gc.ca/cnt/rpt/annrpt/annrpt20142015-eng.aspx>
- Saunders, D. G. (1995). Prediction of wife assault. In Campbell, J. (Ed.), *Assessing Dangerousness: Violence by Sexual Offenders, Batterers, and Child Abusers* (pp. 70-74). Thousand Oaks, CA: Sage Publications.
- Schechter, S. (1982). *Women and male violence: The visions and struggles of the battered women's movement*. Boston, MA: South End Press.
- Schmidt, J. D., & Sherman, L. W. (1996). Does arrest deter domestic violence? In Buzawa, E., and Buzawa, C. (Eds.), *Do Arrests and Restraining Orders Work?* (pp. 43–53). Thousand Oaks, CA: Sage Publications.
- Schneider, E. M. (2000). *Battered women & feminist lawmaking*. New Haven: Yale University Press.
- Schonrock, J., & Fawbuch, J. (2018). *Domestic violence: History of police responses*. <https://www.findlaw.com/family/domestic-violence/domestic-violence-history-of-police-responses.html>

- Schumacher, J. A., Feldbau-Kohn, S., Smith Slep, A. M., & Heyman, R. E. (2001). Risk factors for male-to-female partner physical abuse. *Aggression & Violent Behavior, 6*(1), 281–352. [https://doi.org/10.1016/S1359-1789\(00\)00027-6](https://doi.org/10.1016/S1359-1789(00)00027-6)
- Sharpe, C. E. (2016). *In the wake: On blackness and being* (Ser. Anti-racist collection). Duke University Press. <https://doi.org/10.1515/9780822373452>
- Sheehan, B. E., Murphy, S. B., Moynihan, M. M., Dudley-Fennessey, E., & Stapleton, J. G. (2015). Intimate partner homicide: New insights for understanding lethality and risks. *Violence Against Women, 21*(2), 269-288. <https://doi.org/10.1177/1077801214564687>
- Sherman, L. W. (1992). *Policing Domestic Violence: Experiments and Dilemmas*. New York, NY: Free Press.
- Sherman, L. W. (2000). Domestic violence and restorative justice: Answering key questions. *Virginia Journal of Social Policy and Law, 8*, 263-312.
- Sherman, L. W., & Berk, R. A. (1984). *The Minneapolis domestic violence experiment*. Washington, D.C: Police Foundation.
- Sherman, L. W., Schmidt, J. D., Rogan, D. P., Smith, D. A., Gartin, P. R., Cohn, E. G., Collins, J., & Bacich, A. R. (1992). The variable effects of arrest on criminal careers: The Milwaukee domestic violence experiment. *The Journal of Criminal Law and Criminology, 83*(1), 137–169.
- Sherman, L. W., Smith, D. A., Schmidt, J. D. & Rogan, D. P. (1992). Crime, punishment, and stake in conformity: Legal and informal control of domestic

- violence. *American Sociological Review*, 57(5), 680-690.
<https://www.jstor.org/stable/2095921>
- Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy*, 85(3), 257-268.
<https://doi.org/10.1093/ptj/85.3.257>
- Simpson, S. S., Bouffard, L. A., Garner, J., & Hickman, L. (2006). The influence of legal reform on the probability of arrest in domestic violence cases. *Justice Quarterly*, 23(3), 297. <https://doi.org/10.1080/07418820600869087>
- Smith, A. (2001). Domestic violence laws: The voices of battered women. *Violence and Victims*, 16(1), 91–111. <https://doi.org/10.1891/0886-6708.16.1.91>
- Sparks, A. (1997). Feminists negotiate the executive branch: The policing of male violence. In C. R. Davies (Ed.), *Feminists negotiate the state: The politics of domestic violence* (pp. 35-52). Lanham, MD: University Press of America.
- Stark, E. (2004). Insults, injury, and injustice: Rethinking state intervention in domestic violence cases. *Violence against Women*, 10(11), 1302–1330.
<https://doi.org/10.1177/1077801204269348>
- Stark, E. (2009). *Coercive control: How men entrap women in personal life*. New York, NY: Oxford University Press.
- Statistics Canada. (2016). Family violence in Canada: A statistical profile, 2014. *The Daily*. <http://www.statcan.gc.ca/daily-quotidien/160121/dq160121b-eng.htm>
- Statistics Canada. (2020a). *Community safety and crime*. (Catalogue no. 11-631-X)
<https://www150.statcan.gc.ca/n1/pub/11-631-x/2020004/s10-eng.htm>

- Statistics Canada (2020b). Experiences of discrimination during the COVID-19 pandemic. *The Daily*. Retrieved from <https://www150.statcan.gc.ca/n1/daily-quotidien/200917/dq200917a-eng.htm>
- Steadman, H. J., & Cocozza, J. J. (1974). *Careers of the criminally insane: Excessive social control of deviance*. Lexington, MA: Lexington Books.
- Stewart, A. (2001). Policing domestic violence: An overview of emerging issues. *Police Practice and Research: An International Journal*, 2(4), 447–460. https://research-repository.griffith.edu.au/bitstream/handle/10072/3851/16496_1.pdf?sequence=1&isAllowed=y
- Storey, J. E., Gibas, A. L., Reeves, K. A., & Hart, S. D. (2011). Evaluation of violence risk (threat) assessment training for police and other criminal justice professionals. *Criminal Justice and Behavior*, 38(6), 554 – 563. <https://doi.org/10.1177/0093854811403123>
- Storey, J. E., Kropp, P. R., Hart, S. D., Belfrage, H., & Strand, S. (2014). Assessment and management of risk for intimate partner violence by police officers using the Brief Spousal Assault Form for the Evaluation of Risk. *Criminal Justice and Behavior*, 41(2), 256–271. <https://doi.org/10.1177/0093854813503960>
- Storey, J. E., & Strand, S. (2013). Assessing violence risk among female IPV perpetrators: An examination of the B-SAFER. *Journal of Aggression, Maltreatment & Trauma*, 22(9), 964-980. <https://doi.org/10.1080/10926771.2013.835015>

- Storey, J. E., & Strand, S. (2017). The influence of victim vulnerability and gender on police officers' assessment of intimate partner violence risk. *Journal of Family Violence, 32*(1), 125-134. <https://doi.org/10.1007/s10896-016-9905-0>
- Strauchler, O., McCloskey, K., Malloy, K., Sitaker, M., Grigsby, N., & Gillig, P. (2004). Humiliation, manipulation, and control: Evidence of centrality in domestic violence against an adult partner. *Journal of Family Violence, 19*, 339–346. <https://doi.org/10.1007/s10896-004-0679-4>
- Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest, 1*(1), 1-26. <https://doi.org/10.1111%2F1529-1006.001>
- Tabachnick, B.G. & Fidell, L.S. (2013). *Using Multivariate Statistics*. Boston, MA: Pearson Education.
- Telep, C. W., & Lum, C. (2014). The receptivity of officers to empirical research and evidence-based policing: An examination of survey data from three agencies. *Police Quarterly, 17*(4), 359–385. <https://doi.org/10.1177/1098611114548099>
- Thijssen, J., & de Ruiter, C. (2011). Identifying subtypes of spousal assaulters using the B-SAFER. *Journal of Interpersonal Violence, 26*(7), 1307-1321. <https://doi.org/10.1177/0886260510369129>
- Thornberry, T. P., & Jacoby, J. E. (1979). *The criminally insane: A community follow-up of mentally ill offenders*. Chicago, IL: University of Chicago Press.
- Tjaden, P., & Thoennes, N. (2000). *Extent, Nature, and Consequences of Intimate Partner Violence: Findings from the National Violence Against Women Survey*.

Washington, D.C.: U.S. Department of Justice, Office of Justice Programs.

<https://www.ncjrs.gov/pdffiles1/nij/181867.pdf>

Trujillo, M. P., & Ross, S. (2008). Police response to domestic violence: Making decisions about risk and risk management. *Journal of Interpersonal Violence*, 23(4), 454-473. <https://doi.org/10.1177/0886260507312943>

Tweed, R. G., & Dutton, D. G. (1998). A comparison of impulsive and instrumental subgroups of batterers. *Violence and Victims*, 13(3), 217-230. <https://doi.org/10.1891/0886-6708.13.3.217>

VanderPut, C. E., Gubbels, J., & Assink, M. (2019). Predicting domestic violence: A meta-analysis on the predictive validity of risk assessment tools. *Aggression and Violent Behavior*, 47, 100–116.

Varcoe, C., Browne, A. J., Ford-Gilboe, M., Dion Stout, M., McKenzie, H., Price, R., Bungay, V., Smye, V., Inyallie, J., Day, L., Khan, K., Heino, A., & Merritt-Gray, M. (2017). Reclaiming our spirits: Development and pilot testing of a health promotion intervention for Indigenous women who have experienced intimate partner violence. *Research in Nursing & Health*, 40(3), 237–254. <https://doi.org/10.1002/nur.21795>

Varcoe, C., Hankivsky, O., Ford-Gilboe, M., Wuest, J., Wilk, P., Hammerton, J., & Campbell, J. (2011). Attributing selected costs to intimate partner violence in a sample of women who have left abusive partners: A social determinants of health approach. *Canadian Public Policy*, 37(3), 359-380. <https://doi.org/10.3138/cpp.37.3.359>

- Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine, 37*, 360-363. http://web2.cs.columbia.edu/~julia/courses/CS6998/Interrater_agreement.Kappa_statistic.pdf
- Viljoen, J. L., Beneteau, J. L., Brodersen, E., Gulbransen, E., Desmarais, S. L., Nicholls, T. L., & Cruise, K. R. (2012). Assessment of multiple risk outcomes, strengths, and change with the START:AV: A short-term prospective study with adolescent offenders. *International Journal of Forensic Mental Health, 11*(3), 165-180. <https://doi.org/10.1080/14999013.2012.737407>
- Viljoen, J. L., Cochrane, D. M., & Jonnson, M. R. (2018). Do risk assessment tools help manage and reduce risk of violence and reoffending? A systematic review. *Law and Human Behavior, 42*(3), 181–214. <https://doi.org/10.1037/lhb0000280>
- Viljoen, J. L., Gray, A. L., Shaffer, C., Bhanwer, A., Tafreshi, D., & Douglas, K. S. (2016). Does Reassessment of Risk Improve Predictions? A Framework and Examination of the SAVRY and YLS/CMI. *Psychological Assessment, 29*(9), 1096–1110. <https://doi.org/10.1037/pas0000402>
- Viljoen, J. V., Gray, A. L., Shaffer, C., Latzman, N. W., Scalora, M. J., & Ullman, D. (2017). Changes in J-SOAP-II and SAVRY scores over the course of residential, cognitive-behavioral treatment for adolescent sexual offending. *Sexual Abuse: Journal of Research and Treatment, 29*(4), 342-374. <https://doi.org/10.1177/1079063215595404>
- 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(4), 549–576. <https://doi.org/10.1037/a0027186>
- Vose, B., Smith, P., & Cullen, F. T. (2013). Predictive validity and the impact of change in total LSI-R score on recidivism. *Criminal Justice and Behavior*, 40(12), 1383-1396. <https://psycnet.apa.org/doi/10.1177/0093854813508916>
- Waltz, J., Babcock, J. C., Jacobsen, N. S., & Gottman, J. M. (2000). Testing a typology of batterers. *Journal of Consulting and Clinical Psychology*, 68(4), 658-669. <https://doi.org/10.1037/0022-006X.68.4.658>
- Ward, T., Melsner, J., & Yaes, P. M. (2007). Reconstructing the Risk–Need–Responsivity model: A theoretical elaboration and evaluation. *Aggression and Violent Behaviour*, 12(1), 208-228. <http://dx.doi.org/10.1016/j.avb.2006.07.001>
- Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997) *HCR-20: Assessing risk for violence, version 2*. Burnaby, BC: Mental Health, Law and Policy Institute, Simon Fraser University.
- Weisz, A. N., Tolman, R. M., & Saunders, D. G. (2000). Assessing the risk of severe domestic violence: The importance of survivors' predictions. *Journal of Interpersonal Violence*, 15(1), 75-90. <https://doi.org/10.1177/088626000015001006>
- West, C. M., Kantor, G. K., & Jasinski, J. L. (1998). Sociodemographic predictors and cultural barriers to help-seeking behavior by Latina and Anglo American battered women. *Violence and Victims*, 13(4), 361–375.

- Whitaker, D. J., & Lutzker, J. R. (2009). *Preventing partner violence: Research and evidence-based intervention strategies*. Washington, D.C.: American Psychological Association. <https://doi.org/10.1037/11873-000>
- Whitehead, J., Dawson, M., & Hotton, T. (2021). Same-sex intimate partner violence in Canada: Prevalence, characteristics, and types of incidents reported to police services. *Journal of Interpersonal Violence, 36*(23-24), 10959-10988. <https://doi.org/10.1177/0886260519897342>
- Williams, K. R., & Stansfield, R. (2017). Disentangling the risk assessment and intimate partner violence relation: Estimating mediating and moderating effects. *Law and Human Behavior, 41*(4), 344-353. <https://doi.org/10.1037/lhb0000249>
- Wilson, C. M., Desmarais, S. L., Nicholls, T. L., Hart, S. D., & Brink, J. (2013). Predictive validity of dynamic factors: Assessing violence risk in forensic psychiatric inpatients. *Law and Human Behavior, 37*(6), 377-388. <https://psycnet.apa.org/doi/10.1037/lhb0000025>
- Wilson, M., Johnson, H., & Daly, M. (1995). Lethal and non-lethal violence against wives. *Canadian Journal of Criminology, 37*(3), 331-362. <http://dx.doi.org/10.4135/9781452243306.n8>
- Women's Equality Branch. (2014). Woman victims of abuse protocols. *Government of New Brunswick*. <https://www2.gnb.ca/content/dam/gnb/Departments/ecobce/WEB-EDF/Violence/PDF/en/WVAP2014-E-jan2016.pdf>
- Wood, J., Sorg, E. T., Groff, E. R., Ratcliffe, J. H., & Taylor, C. J. (2014). Cops as treatment providers: Realities and ironies of police work in a foot patrol

experiment. *Policing and Society*, 24(3), 362–379.

<https://doi.org/10.1080/10439463.2013.784292>

Wortley, S. (2006). *Police use of force in Ontario: An examination of data from the special investigations unit, final report*. Toronto: African Canadian Legal Clinic.
http://www.archives.gov.on.ca/en/e_records/ipperwash/policy_part/projects/pdf/AfricanCanadianClinicIpperwashProject_SIUStudybyScotWortley.pdf

Wortley, S., & Owusu-Bempah, A. (2009). Unequal before the law: Immigrant and racial minority perceptions of the Canadian criminal justice system. *Journal of International Migration and Integration*, 10(4), 447–473.
<https://doi.org/10.1007/s12134-009-0108-x>

Yesberg, J. A., & Polaschek, D. L. L. (2015). Assessing dynamic risk and protective factors in the community: Examining the validity of the Dynamic Risk Assessment for Offender Re-entry. *Psychology, Crime & Law*, 21(1), 80-99.
<https://doi.org/10.1080/1068316X.2014.935775>

Yoshihama, M., & Horrocks, J. (2003). The relationship between intimate partner violence and PTSD? An application of cox regression with time-varying covariates. *Journal of Traumatic Stress*, 16(4), 371-380.
<https://doi.org/10.1023/a:1024418119254>

Yu, R., Nevado-Holgado, A. J., Molero, Y., D'Onofrio, B. M., Larsson, H., Howard, L. M., & Fazel, S. (2019). Mental disorders and intimate partner violence perpetrated by men towards women: A Swedish population-based longitudinal study. *Plos Medicine*, 16(12), 1-19.
<https://doi.org/10.1371/journal.pmed.1002995>

- Zahra, D., & Hedge, C. (2010). The reliable change index: Why isn't it more popular in academic psychology? *PsyPag Quarterly*, 76(6), 14-19.
<https://doi.org/10.1111/medu.13059>
- Zhang, T., Hoddenbagh, J., McDonald, S., & Scrim, K. (2012) *An estimation of the economic impact of spousal violence in Canada, 2009*. Ottawa, ON: Department of Justice Canada. https://www.justice.gc.ca/eng/rp-pr/cj-jp/fv-vf/rr12_7/rr12_7.pdf
- Zinger, I. (2021). *Annual report: 2020 – 2021*. The Office of the Correctional Investigator. <https://www.oci-bec.gc.ca/cnt/rpt/pdf/annrpt/annrpt20202021-eng.pdf>

Table 1

Results of Chi-square Test and Descriptive Statistics for Perpetrator Characteristics, Overall and as a Function of Perpetrator Gender

Variable	% Males (<i>n</i> = 209)	% Females (<i>n</i> = 100)	χ^2	df	% Overall (<i>N</i> = 309)
Pre-index IPV offense(s)	42.1 (88)	25.0 (25)	8.53**	1	36.6 (113)
Pre-index violent offense(s)	72.7 (152)	52.0 (52)	12.95***	1	66 (204)
Pre-index general offense(s)	79.4 (166)	43.0 (43)	41***	1	67.6 (209)
History of IPV victimization	8.6 (18)	37.0 (37)	37.25***	1	17.8 (55)
Recidivism within 1 year of index	35.9 (75)	20.0 (20)	8.02**	1	30.4 (95)
Substance use at index	45.5 (95)	35.0 (35)	3.03	1	42.1 (130)
Weapon use at index	13.9 (29)	11.0 (11)	.50	1	12.9 (40)

Note. IPV = Intimate partner violence. ****p* < .001.

Table 2

Results of Chi-square Test and Descriptive Statistics for Victim Characteristics, as a Function of Victim Gender

Variable	% Males (<i>n</i> = 94)	% Females (<i>n</i> = 215)	χ^2	df	% Overall (<i>N</i> = 309)
Did not cooperate with police investigation	64.9 (26)	47.4 (102)	8.26*	1	52.8 (163)
Had a visible injury	48.9 (46)	48.4 (104)	.01	1	48.5 (150)
Substance use at index	23.4 (22)	28.5 (61)	.86	1	26.9 (83)

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

Table 3

Results of Chi-square Test and Descriptive Statistics for Type of Abuse Committed
During the Index Offense by Gender

Type of IPV (Yes or No):	% Males (<i>n</i> = 209)	% Females (<i>n</i> = 100)	χ^2	df	% Overall (<i>N</i> = 309)
Physical Abuse	82.8 (173)	78.0 (78)	1.01	1	81.2 (251)
Emotional Abuse	56.9 (119)	53.0 (53)	.43	1	55.7 (172)
Property Abuse	29.7 (62)	34.0 (34)	.59	1	31.1 (96)
Stalking/Harassment	17.2 (36)	15.0 (15)	.24	1	16.5 (51)
Financial Abuse	4.3 (9)	2.0 (2)	1.05	1	3.6 (11)

Note. IPV = Intimate partner violence. **p* < .05. ***p* < .01. ****p* < .001.

Table 4

Descriptive Statistics and Results of ANOVAs Comparing Males and Female Suspects on Number of Different IPV Actions Perpetrated Within Each Abuse Category

Variable	Male		Female		<i>F</i>	<i>df</i>	Overall	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			<i>M</i>	<i>SD</i>
Physical Abuse	2.02	1.58	1.37	1.22	13.11***	1, 305	1.80	1.50
Emotional Abuse	1.17	1.47	.94	1.20	1.83	1, 305	1.09	1.39
Property Abuse	.33	.54	.36	.52	.26	1, 305	.34	.53
Stalking/Harassment	.31	.78	.23	.60	.87	1, 305	.29	.73
Financial Abuse	.05	.24	.02	.14	1.19	1, 305	.04	.21

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

Table 5

*Results of Chi-square Test and Descriptive Statistics for Reason for the Altercation
During the Index Offense by Gender*

Reason for Altercation:	% Males (<i>n</i> = 209)	% Females (<i>n</i> = 100)	χ^2	df	% Overall (<i>N</i> = 309)
Jealousy	31.6 (66)	23.0 (23)	2.43	1	28.8 (89)
Unknown/other	27.3 (57)	29.0 (29)	.10	1	27.8 (86)
Children	10.5 (22)	25.0 (25)	10.99***	1	15.2 (47)
Breaking up	12.9 (27)	9.0 (9)	1.01	1	11.7 (36)
Substance use	7.2 (15)	6.0 (6)	.15	1	6.8 (21)
Money	8.1 (17)	2.0 (2)	4.41*	1	6.1 (19)
Possessions/ownership	5.3 (11)	6.0 (6)	.07	1	5.5 (17)
Infidelity	3.8 (8)	4.0 (4)	.01	1	3.9 (12)
Sleep disrupted	2.4 (5)	3.0 (3)	.10	1	2.6 (8)
Sex	1.9 (4)	1.0 (1)	.36	1	1.6 (5)

Note. **p* < .05. ***p* < .01. ****p* < .001.

Table 6

Frequency and Percent of IPV Files with Police-Scored ODARA at the Index and Subsequent Recidivism within 12-Month Follow Up Period

Variable	ODARA Noted in File		ODARA Completed in Full	
	Frequency	% (N)	Frequency	% (N)
Index	118	61.8 (191)	94	79.7 (118)
Recidivism 1	16	20.3 (79)	8	50.0 (16)
Recidivism 2	4	14.3 (28)	2	50.0 (4)
Recidivism 3	1	9.1 (11)	1	100.0 (1)
Recidivism 4	3	42.9 (7)	2	66.7 (3)
Recidivism 5	0	0 (4)	--	--
Recidivism 6	0	0 (3)	--	--
Recidivism 7	0	0 (1)	--	--

Note. IPV = Intimate partner violence. Includes only files whose Index dates 2015 or later.

Table 7

Interrater Reliability Estimates between Police Officers and Researchers for ODARA Total Scale Scores (ICC) and Item Scores (Kappa)

Variable	ICC _{2,2}	Kappa
Sum of ODARA Items	.61	-
Item 1. Has a prior domestic incident	-	.26
Item 2. Has a prior nondomestic incident	-	.25
Item 3. Has a prior sentence of 30 days or more	-	.23
Item 4. Has failed on prior conditional release	-	.48
Item 5. Made threats to harm or kill during the index	-	.58
Item 6. Confinement of the partner at the index event	-	.63*
Item 7. Victim fears repetition of violence	-	.34
Item 8. Victim and/or offender have more than one child	-	.67*
Item 9. Victim has biological child from previous partner	-	.61*
Item 10. Offender is violent outside of this relationship	-	.27
Item 11. Indicator of problematic substance use	-	.49
Item 12. Offender has assaulted victim when pregnant	-	.79*
Item 13. Victim has at least one barrier to support	-	.09

Note. * = acceptable agreement. Includes only post-2015 files. ODARA items from Hilton et al. (2004). ICC_{2,1} = two-way, random effects model. ICC \geq .70 is acceptable. Kappas of .41 to .60 = moderate agreement, .61 to .80 = substantial agreement, and \geq .81 = excellent agreement.

Table 8

Summary of Logistic Regression Analysis for Suspect and Contextual Details in Predicting Police Use of ODARA at Index (N = 191)

	<i>B</i>	<i>SE</i>	Wald	<i>df</i>	Odds Ratio
Suspect					
Gender	-1.00	0.40	6.15*	1	0.37
Substance Use	0.70	0.34	4.30*	1	2.10
Contextual					
Weapon Used	-0.78	0.55	1.99	1	0.46
Currently in Relationship	0.61	0.40	2.33	1	1.84
Living Arrangement					
Living Together (reference)			0.37	2	
Living Separately	0.09	0.51	0.04	1	1.11
Unknown	-0.16	0.55	0.08	1	0.85
Visible Injury	-0.57	0.33	2.94	1	0.57
Victim Cooperation					
Victim Cooperated (reference)			0.42	3	
Did Not Cooperate	0.17	0.56	0.09	1	1.19
Not applicable	0.26	0.55	0.22	1	1.29
Unknown	0.85	1.58	0.29	1	2.33
<i>R</i> ² (Nagelkerke method)	.16				

Note. Gender is for females compared to males. **p* < .05. ***p* < .01. ****p* < .001.

Table 9

Predictive Validity of ODARA Total Score for Predicting IPV Recidivism as Measured by ROC Analyses

	<i>AUC (95% CI)</i>		
	Full Post-2015 Sample (<i>N</i> = 181)	Females (<i>n</i> = 35)	Males (<i>n</i> = 146)
Police-rated ODARA total score ^a	.58 (.47, .69)	.71 (.47, .96)	.53 (.42, .65)
Researcher-rated ODARA total score	.69 (.60, .76)	.79 (.64, 1.00)	.64 (.60, .74)

Note. ODARA = Ontario Domestic Assault Risk Assessment (Hilton et al., 2004); IPV = intimate partner violence; ROC = Receiver Operator Characteristic; AUC = Area Under the Curve. ^a Sample limited to only post-2015 files for which there was a police-scored ODARA, resulting in *N* = 114 in total, *n* = 15 for females, and *n* = 99 for males. Also represented in Figure 1 (researcher-scored) and Figure 2 (police-scored).

Table 10

Number of Actions Taken by Police when Responding to IPV Calls as a Function of Suspect Gender: Results of ANOVAs and Descriptive Statistics

Variable	Male		Female		<i>F</i>	<i>df</i>	Overall	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			<i>M</i>	<i>SD</i>
Total Actions	9.52	3.70	6.97	2.86	14.98***	1, 189	9.04	3.69
Official Criminal Justice Actions	6.07	2.03	5.2	2.23	4.91**	1, 189	5.91	2.09
Victim-Directed Discretionary Actions	3.30	2.49	1.33	1.71	20.08***	1, 189	2.93	2.48
Suspect-Directed Discretionary Actions	.10	.32	.33	.59	11.22***	1, 189	.14	.39

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

Table 11*Results of Chi-square Test for Risk Management Strategies by Police-Rated Risk**Classification (Post 2015 files with Police-Scored ODARAs)*

Variable	Low % (n= 22)	Mod. % (n =53)	High % (n = 39)	χ^2	Overall % (N= 114)
Official CJ Actions					
Looked up record	86.4 (19)	100 (53)	97.4 (38)	8.70*	96.5 (110)
Interviewed 1 party	36.4 (8)	56.6 (30)	51.3 (20)	2.55	50.9 (58)
Interviewed both	72.7 (16)	66 (35)	64.1 (25)	0.49	66.7 (76)
Arrested suspect	68.2 (15)	81.1 (43)	97.4 (38)	9.76**	84.2 (96)
Rec. charges	63.6 (14)	83 (44)	92.3 (36)	8.02*	82.5 (94)
Rec. pretrial detention	9.1 (2)	41.5 (22)	59.0 (23)	14.45***	41.2 (47)
Rec. conditions	63.6 (14)	66 (35)	57.9 (38)	0.64	62.8 (71)
Called Social Services	22.7 (5)	28.3 (15)	46.2 (18)	4.60	33.3 (38)
Unofficial CJ Actions					
Explain CJS to vic.	50 (11)	54.7 (29)	52.6 (38)	0.14	53.1 (60)
Victim Services referral	45.5 (10)	52.8 (28)	56.4 (22)	0.68	52.6 (60)
Other victim referral	27.3 (6)	32.1 (17)	38.5 (15)	0.86	33.3 (38)
Safety plan with vic.	18.2 (4)	34 (18)	43.6 (17)	4.04	34.2 (39)
IPV educated vic.	19 (4)	26.4 (14)	28.2 (11)	0.63	25.7 (29)
Other suspect referral	4.5 (1)	0 (0)	0 (0)	4.22	0.9 (1)

Note. CJ = Criminal Justice; Rec. = recommended to the Crown attorney; Explain CJS to vic. = explained to victim how the investigation or other criminal justice system proceedings would go; IPV educated vic. = explained seriousness of Intimate Partner Violence to victim and/or shared results of ODARA; * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 12

*Kendall's Tau-b Correlations Between Dynamic Risk Factors and Dichotomous IPV
Recidivism Outcome*

Dynamic Risk Factor	τ_b					
	Any IPV	Physical	Emotional	Property	Stalking/ Harassment	Financial
Victim Fear	.07	.24**	.11	-.05	.26**	.06
Antisocial Attitudes	.13*	-.02	.02	.04	.01	.03
Suicide/Self-Harm	.08	-.13	-.06	.10	.05	.14
Jealousy/ Control	.10	-.08	.11	.07	.07	-.13
Employment/ Financial Problems	.08	.05	.03	-.11	.20	.10
Alcohol/Drug Use	-.04	.10	-.003	-.12	-.06	.10
Relationship Problems	.06	-.07	-.18	.11	.09	.04

Note. * $p < .05$. ** $p < .01$. *** $p < .001$. τ_b = Kendall's Tau-b. IPV = Intimate Partner Violence.

Table 13*Mean Survival Time for Three Levels of Dynamic Risk Factors*

Dynamic Risk Factor	Mean (SE) Time in Days		
	No Evidence	Partial Evidence	Evidence
Victim Fear	296.73 (9.04)	277.42 (27.44)	274.20 (15.41)
Antisocial Attitudes	308.80 (10.20) ^a	291.91 (18.97)	266.65 (12.84) ^c
Suicide/Self-Harm	291.43 (7.93)	254.88 (40.21)	267.75 (30.89)
Jealousy/Controlling	293.19 (9.95)	318.55 (16.06)	266.53 (15.06)
Employment/Financial Problems	293.84 (8.21)	277.30 (26.76)	253.36 (27.26)
Alcohol/Drug Use	287.24 (10.12)	288.40 (38.42)	302.59 (11.32)
Relationship Problems	293.49 (16.03)	325.88 (15.96)	279.71 (7.60)

Note. Significant differences between dynamic risk levels at the $p < .05$ level denoted by use of "a", "b", "c".

Table 14*Reliable Change Index (RCI) Summary Statistics for Dynamic Risk Factors*

Measure	Reliable Decrease		No Reliable Change		Reliable Increase	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Victim Fear	24	27.6	54	62.1	9	10.3
Antisocial Attitudes	29	33.3	36	41.4	22	25.2
Suicide/Self-Harm	12	13.8	73	83.9	2	2.2
Jealousy/Controlling	25	28.7	50	57.5	12	13.8
Employment/Financial Problems	16	18.3	65	74.7	6	6.9
Alcohol/Drug Use	20	23.0	57	65.5	10	11.5
Relationship Problems	19	21.8	50	57.5	18	20.6
Total DRF Score	58	66.7	8	9.2	21	24.1

Note. $N = 87$. DRF = Dynamic Risk Factor.

Table 15

Predictive Validity of 'Antisocial Attitudes' Dynamic Risk Item for Predicting IPV Recidivism as Measured by ROC Analyses, Compared by Suspect Gender

Follow-Up Period	AUC [95% CI]			% CI Overlap
	Full Sample (N = 298)	Females (n = 99)	Males (n = 199)	
1 month	.588 [.490, .685]	.518 [.317, .719]	.592 [.481, .703]	55.22
3 months	.598 [.514, .683]	.542 [.368, .716]	.600 [.503, .696]	55.46
6 months	.594 [.519, .669]	.536 [.381, .690]	.594 [.507, .681]	56.31
12 months	.578 [.507, .649]	.548 [.404, .692]	.568 [.484, .651]	57.99

Note. IPV = Intimate partner violence; ROC = Receiver operating curve; AUC = Area under the curve; CI = Confidence intervals; % CI Overlap = Percent of overlap between confidence intervals for males and females.

Table 16

Predictive Validity of 'Antisocial Attitudes' Dynamic Risk Item for Predicting IPV Recidivism as Measured by ROC Analyses, Compared to Researcher-Completed ODARA Total Score

Follow-Up Period	AUC [95% CI]		% CI Overlap
	Antisocial Attitudes	ODARA	
1 month	.588 [.490, .685]	.667 [.572, .762]	27.35
3 months	.598 [.514, .683]	.675 [.592, .757]	25.42
6 months	.594 [.519, .669]	.664 [.591, .737]	26.35
12 months	.578 [.507, .649]	.657 [.588, .726]	21.79

Note. IPV = Intimate partner violence; ROC = Receiver operating curve; AUC = Area under the curve; CI = Confidence intervals; % CI Overlap = Percent of overlap between confidence intervals for ODARA and Total Dynamic Risk Score. *N* = 298.

Table 17

Incremental Validity of Adding the ‘Antisocial Attitudes’ Dynamic Risk Item to the ODARA for Predicting IPV Recidivism as Measured by ROC Analyses

Follow-Up Period	AUC [95% CI]		% CI Overlap
	ODARA	ODARA + AA	
1 month	.667 [.572, .762]	.678 [.582, .774]	47.12
3 months	.675 [.592, .757]	.683 [.601, .766]	47.27
6 months	.664 [.591, .737]	.673 [.599, .746]	47.10
12 months	.657 [.588, .726]	.662 [.593, .731]	48.19

Note. IPV = Intimate partner violence; ROC = Receiver operating curve; AUC = Area under the curve; CI = Confidence intervals; ODARA + AA = ODARA plus dichotomous ‘antisocial attitudes item; % CI Overlap = Percent of overlap between confidence intervals for ODARA and Total Dynamic Risk Score. *N* = 298.

Table 18

Direction of Change in ODARA Risk Score from Index to First Recidivism Event as a Function of Risk Classification at Index

ODARA Risk Classification	Reliable Decrease		No Reliable Change		Reliable Increase	
	Frequency	χ^2	Frequency	χ^2	Frequency	χ^2
Low Risk	0 (-1.03)	1.06	0 (-2.00)	4.00	6 (2.68)	7.18*
Moderate Risk	1 (-2.17)	1.41	5 (-3.15)	9.92*	25 (4.59)	21.07**
High Risk	12 (-2.61)	6.81	30 (-4.05)	16.40*	12 (-5.78)	33.41**

Note. $N = 87$. χ^2 = chi squared value. Adjusted residuals appear in parentheses below observed frequencies. A Bonferroni correction was applied, resulting in * = $p < .0056$, and ** = $p < .0011$.

Table 19

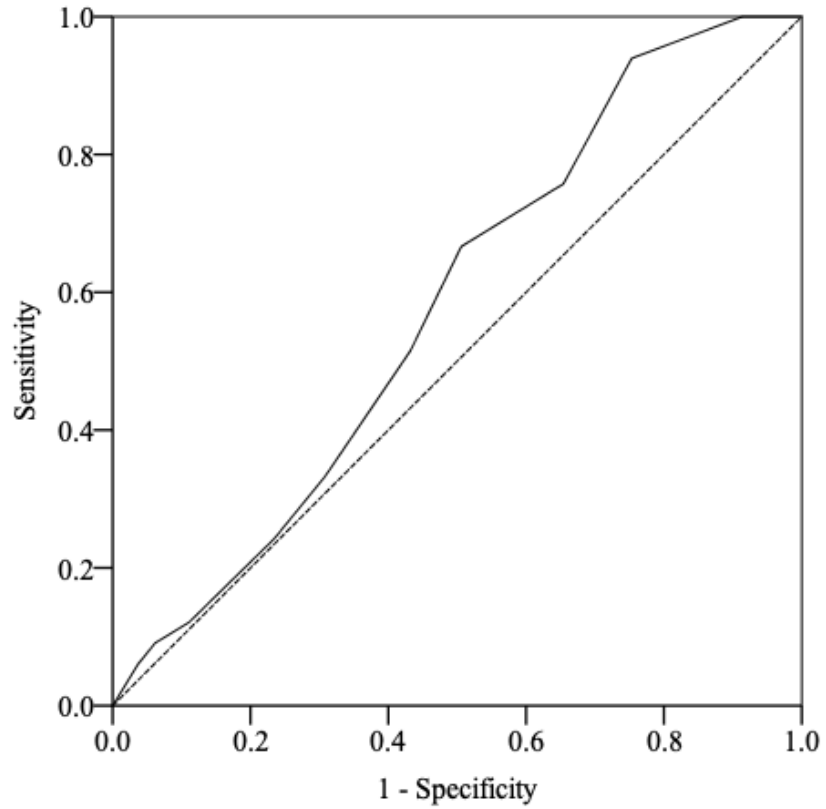
Direction of Change in ODARA + AA Score from Index to First Recidivism Event as a Function of Risk Classification at Index

ODARA Risk Classification	Reliable Decrease		No Reliable Change		Reliable Increase	
	Frequency	χ^2	Frequency	χ^2	Frequency	χ^2
Low Risk	0 (-1.69)	2.86	1 (-.50)	0.25	5 (2.01)	4.04
Moderate Risk	3 (-3.13)	9.80*	7 (-.43)	0.18	21 (3.29)	10.82**
High Risk	25 (3.88)	15.05**	15 (.66)	0.44	14 (-4.19)	17.56**

Note. $N = 87$. χ^2 = chi squared value. Adjusted residuals appear in parentheses below observed frequencies. A Bonferroni correction was applied, resulting in * = $p < .0056$, and ** = $p < .0011$.

Figure 1

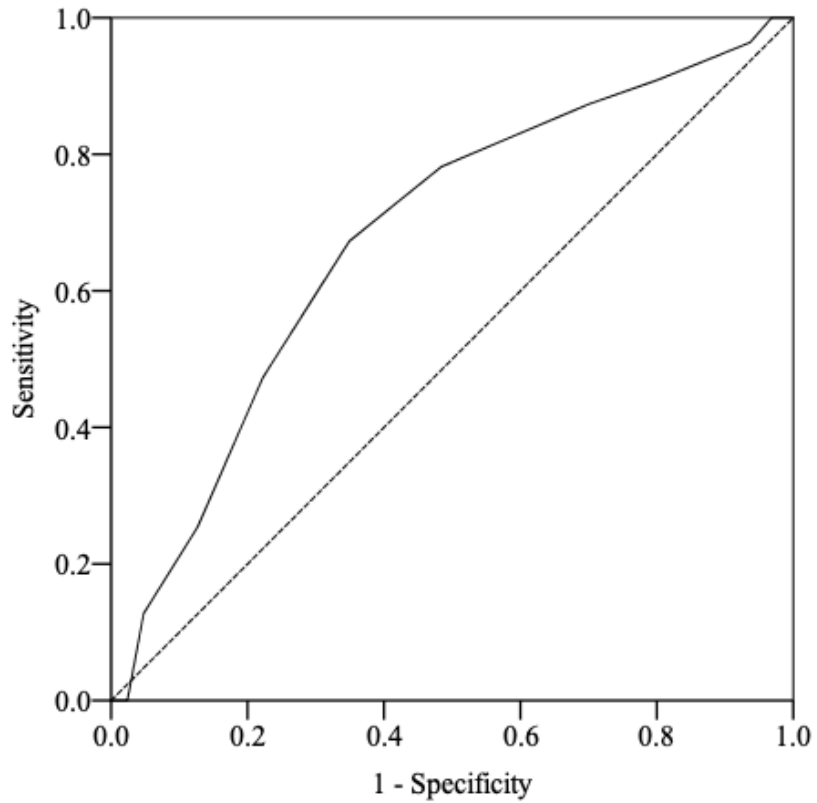
ROC Curve for Police-Rated ODARA Score and Recidivism



Note. Using the post-2015 sample for which there was a police-scored ODARA available ($N=114$).

Figure 2

ROC Curve for Researcher-Rated ODARA Score and Recidivism



Note. Using the post-2015 sample for which there was a police-scored ODARA available ($N=114$).

Appendix A

ODARA Item Scoring and Types of Risk Factors

*Items Coded in the Ontario Domestic Assault Risk Assessment (ODARA),
Corresponding Operational Definitions, Risk Factor Type, and Scoring Protocol*

Variable and Definition	Type	Score
1. Has a <i>prior domestic incident</i> in police report or criminal record against any current or former intimate partner, or physical assault against their children.	Static	0 or 1
2. Has a <i>prior nondomestic incident</i> in police or criminal record (i.e., any previous act of physical violence inflicted on anyone other than current/former intimate partner or their children).	Static	0 or 1
3. Has a <i>prior sentence of 30 days or more</i> (including training school, jail, or prison), even if it was served intermittently. Do not count time served in custody while waiting for court.	Static	0 or 1
4. Has <i>failed on prior conditional release</i> (bail, parole, probation, no-contact order), whether civil or criminal. Include any known violation, even if it was not reported to the police or charged. Include any breaches that occur any time before or up to the time of the IPV.	Static	0 or 1
5. Made <i>threats to harm or kill during the index event</i> (e.g., threatens victim, police, witness, child). Only include threats of physical harm/death against people. Do not include threats of non-physical harm (e.g., vandalism, emotional harm, financial harm, legal action, custody dispute). Only include gestures that are commonly/objectively seen as threats to harm/kill (i.e., it is not enough that the victim felt threatened by a gesture, it must be a gesture that anyone would objectively see as threatening). Only include threats that were a part of the index IPV event (i.e., do not include	Dynamic	0 or 1

threats made previously or after the index event). Threats made as the suspect is being arrested, interviewed, or detained immediately following the IPV index event can be counted.

- | | | |
|---|---------|--------|
| <p>6. <i>Confinement of the partner at the index</i> event, such that the suspect physically prevents (or attempts to prevent) them leaving the scene of the incident (including locking doors/windows, standing between victim and the escape route, taking the car keys, damaging the escape vehicle, holding onto the victim/their hair/clothes). Do not include threats to harm (etc.) if the victim leaves. Do not include restricting access to phones. Do not include confinement prior to the index event.</p> | Dynamic | 0 or 1 |
| <p>7. <i>Victim fears repetition of violence</i> and explicitly makes a statement that indicates concern, fear, worry, or certainty about future IPV. Include any expression of fear that occurred after the IPV, even if the victim later changes their mind and claims they are not afraid. Do not include statements of fear for safety during the index IPV (e.g., “when his hands were around my throat, I was afraid he was going to kill me). Do not include statements of concern, fear, worry, or certainty about future IPV that were expressed prior to the index event. Do not include inferences about concern based on the victim’s reaction to the IPV or the protective action they take afterwards (e.g., do not include if the victim decides to move so the suspect doesn’t know where they live).</p> | Dynamic | 0 or 1 |
| <p>8. Victim and/or offender have <i>more than one child</i>, adopted or biological. Include adult children and children who live elsewhere.</p> | Static | 0 or 1 |

9. <i>Victim has biological child from previous partner.</i> Include only biological children. Include adult children and children who live elsewhere.	Static	0 or 1
10. Offender is <i>violent outside of this relationship</i> , regardless of if it was reported to the police or if they were charged. Do not include violence against any current or former intimate partner or their children.	Static	0 or 1
11. Two or more indicators of <i>problematic substance use</i> . Include:	Static or Dynamic	0 or 1
<ul style="list-style-type: none"> • Consumed alcohol immediately before or during IPV index • Consumed drugs immediately before or during IPV index • Abused alcohol or used drugs in the days or weeks leading up to the IPV index (alcohol abuse = frequent or excessive use of alcohol, alcohol use to the point of intoxication; drug use = use of any street drug, use of marijuana without a prescription, abuse of medication – whether or not it was prescribed) • Noticeable increased use of alcohol/drugs in days or weeks prior to the IPV index (e.g., went on a bender) • Suspect has ever previously become more angry or violent when using drugs/alcohol • Suspect has ever previously committed a crime while under the influence of drugs/alcohol • Suspect has ever previously had problems in their adult life due to alcohol use (e.g., financial, job, relationship, health problems) • Suspect has ever previously had problems in their adult life due to drug use (e.g., financial, job, relationship, health problems) 		

<p>12. If victim is female: Offender has <i>assaulted victim when she was pregnant</i>, either during the current event or during previous IPV. Include assaults against a pregnant victim even if the offender did not know she was pregnant at the time. If the victim was male, score “?”.</p>	Static	0 or 1
<p>13. <i>Victim has at least one barrier to support</i>, including if they:</p>	Static or Dynamic	0 or 1
<ul style="list-style-type: none"> • have at least 1 child who is <19 years old that they care for • have no phone • have no access to transportation (car, bus pass, etc.) • are geographically isolated (rural dwelling, no close neighbours) • consumed alcohol/drugs before or during the index IPV • have a history of substance use problems, including any of the following: frequent use of alcohol to the point of intoxication; use of any street/illegal drug; use of marijuana without a prescription; abuse of medication 		

Notes. Revised from Hilton et al. (2004). *0 = absent; 1 = present.

Appendix B

Coding Guide for IPV Incidents

ID #: _____

Coder's Initials: _____

Inclusion criteria:

- ✓ File is in English
- ✓ Perpetrator and victim can be unambiguously identified
- ✓ Both parties were over the age of 18 at the time of the index event
- ✓ The victim and perpetrator were either currently or previously in an intimate relationship
- ✓ Police intervention was a result of an issue between the 2 parties in an intimate relationship
- ✓ Responding officer indicated there was some form of domestic dispute, with the perpetrator's aggression intentionally directed at the victim
- ✓ The domestic dispute must meet the operational definition of IPV used in the current study.
 - "An event in which a current or former spouse, common-law partner, or dating partner abuses, attempts to abuse, or threatens to abuse their intimate partner by use of physical, sexual, psychological (including blackmail & stalking), emotional, or financial abuse."
- ✓ No more than 5 items missing from the ODARA

INCIDENT DETAILS:

1. Incident Date (yyyy/mm/dd): _____ Incident Time (call to 911 or police): _____
2. Which offence is this? ___ Index (the first IPV-related incident recorded for the study for this suspect) ___ Recidivism event If recidivism, which # (1, 2, etc.)? _____
3. If this is a repeat incident, is the victim the same as at the index event? ___ Not Applicable ___ Same victim ___ Different victim ___ Unknown

4. What was the resolution of the case?
 No arrest or charges Arrested, no charges (police or crown’s decision?)
 Arrested and charged Arrested, charged and prosecuted (no conviction)
 Arrested, charged, prosecuted and convicted Information not available
 Other: _____

RELATIONSHIP DETAILS:

5. Relationship status (at time of the IPV):
 Currently in intimate relationship Previously in intimate relationship
 Other: _____

6. Living arrangement (at time of the IPV):
 living together (married or common-law)
 living separately and always have (dating)
 living separately, but formerly lived together
 living separately, unknown previous arrangements
 unknown current or former living arrangements
 Other: _____

7. Length of relationship (e.g., if off and on for six years, enter “6 years”)
 If known, specify (months/years): _____ Unknown/no info

8. Length of time since separation (if no longer in a relationship at the time of the IPV):
 Not Applicable
 If known, specify (months/years): _____ Unknown/no info

DEMOGRAPHIC DETAILS – SUSPECT:

9. Age: (in years) DOB: _____ (yyyy/mm/dd) Unknown

10. Gender:
 Male Female Other: _____ Unknown

11. Ethnicity:
 Caucasian Indigenous Other: _____ Unknown

- Current or recent separation/break up/divorce
- Frequent and serious conflict (i.e., mutual conflict, disagreements that lead to arguments/fighting)
- One partner wanting to end relationship (i.e., prior to the current IPV offence)
- Repeated infidelity (on part of suspect, victim, or both)

26. Current or recent (previous 1 month) indications of suspect excessive jealousy or control of victim?

No evidence Partial/inconclusive evidence Evidence of jealousy/control, such as:

- verbal expressions of excessive jealousy
- obsessive checking in on victim
- becomes angry/hostile in response to victim's associations with other people
- accuses victim of being unfaithful without any grounds for the accusation
- controls most or all of victim's daily activities
- attempts to control victims' behaviour (e.g., using victim's children, threats to end relationship, privilege and punishment, blackmail, threats of harm/vandalism/etc.)
- isolates victim from friends and family

27. Evidence of current or recent (previous 1 month) employment and/or financial problem?

No evidence Partial/inconclusive evidence Evidence for problems, such as:

- recently laid off or fired
- long periods of unstable employment
- low or no income
- use of community resources to meet basic needs (e.g., housing, clothing, food)

<ul style="list-style-type: none"> • failure to seek employment despite financial stress
<p>28. Consumption of alcohol or drug use leading up to and/or during IPV incident?</p> <p><input type="checkbox"/> No evidence <input type="checkbox"/> Partial/inconclusive evidence <input type="checkbox"/> Evidence of alcohol or drug use</p>
<p>29. Current or recent (previous 1 month) suicidality/self-harm on part of suspect?</p> <p><input type="checkbox"/> No evidence <input type="checkbox"/> Partial/inconclusive evidence <input type="checkbox"/> Evidence of suicidality, such as:</p> <ul style="list-style-type: none"> • Cutting, head banging, skin burning, etc. • Threats of suicide • Suicidal gestures or attempts • Feigned suicide gestures or attempts
<p>30. Currently or recently (previous 1 month) demonstrates an attitude that is antisocial or condones violence?</p> <p><input type="checkbox"/> No evidence <input type="checkbox"/> Partial/inconclusive evidence <input type="checkbox"/> Evidence for attitude, such as:</p> <ul style="list-style-type: none"> • Violations of court orders (e.g., violation of conditions of parole, bail, probation, or restraining orders, etc.) • Currently denying responsibility for <i>previous</i> IPV or <i>current</i> IPV (e.g., outright denial or minimizing significance of IPV; blaming the victim for IPV; rationalizing or justifying IPV) • Opposed to police presence or involvement (e.g., runs from cops; refuses to let victim call 911; confrontational with cops, uncooperative with cops, resists arrest, assaults a cop) • Procriminal attitudes: Perpetrator sees crime as a legitimate way to get what they want; see crime as a natural part of life/world (e.g., talks positively/openly about crime, threatens to do crimes in order to meet their needs)
<p>31. Victim fears repetition of violence and explicitly makes a statement that indicates concern, fear, worry, or certainty about future IPV.</p> <p><input type="checkbox"/> No evidence <input type="checkbox"/> Partial/inconclusive evidence <input type="checkbox"/> Evidence for victim fear</p>

- Include any expression of fear that occurred after the IPV, even if the victim later changes their mind and claims they are not afraid.
- Do not include statements of fear for safety during the index IPV (e.g., “when his hands were around my throat, I was afraid he was going to kill me).
- Do not include statements of concern, fear, worry, or certainty about future IPV that were expressed prior to the index event.
- If victim fear is not explicitly stated, but rather inferred based on the victim’s reaction to the IPV or the protective action they take afterwards, count as partial evidence.

POLICE ACTION ITEMS:

Instructions: Check off all actions that the police took in relation to the current file.

32. General Criminal Justice Responses:

- Gathered info about suspect’s criminal record (i.e., official background check)
- Completed ODARA
- Interviewed victim
- Interviewed suspect
- Interviewed witnesses
- Removed one party from the scene, but not arrested
- Arrested suspect
- Arrested victim
- Recommended to Crown that charges be laid against suspect
- Recommended to Crown that suspect given pretrial detention
- Recommended suspect for further psychological and/or risk assessment
- Reported event to Department of Social Development (if children witnessed IPV)
- Other: _____

33. Other Suspect-Directed Actions:

- Issued conditions for the suspect to follow (if yes, indicate which):

- no contact with victim
- restricted from victim's residence
- keep the peace & good behaviour
- do not attempt to locate victim
- no possession of weapons
- no drugs and alcohol
- report to bail supervisor
- promise to appear in court
- other: _____
- Conditions were unspecified in file
- Met with suspect post incident to check-in
- Referred suspect to community services (if yes, indicate which services):
 - Dept. of Social Development
 - Mental Health Services
 - Addiction services
 - Anger management group
 - other referral: _____
 - Other suspect-direction action: _____

34. Other Victim-Directed Actions:

- Gave information/instructions to victim about applying for civil procedures (e.g., family court, restraining orders, etc.)
- Gave information about/assisted victim with obtaining legal aid
- Explained how criminal justice system works (e.g., will press charges even if you don't give statement, court appearances, etc.)
- Referred to Victim Services (i.e., internal SJPF service)
- Referred/provided information to victim about appropriate community services (if yes, indicate which):
 - Dept. of Social Development
 - Shelter/safe house
 - Mental Health Services
 - other referral: _____
 - Service referrals unspecified
- Developed or reviewed a safety plan with victim (e.g., what to do if perp returns, etc.)
- Brought or offered to bring victim to safe house/shelter or other safe place to stay the night
- Called victim post-incident to check in (e.g., well-being check, encourage to give a statement)
- Provided education about IPV and risk to victim (e.g., shared results of ODARA, explained the seriousness of IPV)
- Established/confirmed a victim support person

- Established a police contact person for the victim
- Made referral or personally helped improve security at victim's home (e.g., referred a contractor or security company, helped fixed a broken window or install a home alarm system)
- Other: _____

35. Other General Actions:

- Drove by residence post-incident
- Visited residence to check in post-incident
- Other: _____

36. Indicate which of the following types of abuse were perpetrated during the index event. Select all that apply.

<u>Emotional Abuse</u>	<u>Physical Abuse</u>	<u>Other</u>
<ul style="list-style-type: none"> • Constant texting/calling • Name calling • Threats (to harm the victim, vandalism, kill, harm family/friends) • Excessive yelling and screaming at victim • Belittling the victim • Instilling fear in the victim • Humiliates the victim • Controlling behaviours (limit access to money, friends, family; continuous monitoring) • Mind games (Trying to make the victim doubt 	<ul style="list-style-type: none"> • Ripping clothing • Forced entry • Confinement • Throwing objects • Slapping • Hitting/Beating • Kicking/stomping • Throwing to ground • Grabbing • Slam • Biting • Hitting with object • Hair pulling • Hit head off things • Pushing • Choking 	<p><u>Stalking/Harassment</u></p> <ul style="list-style-type: none"> • Repeated, unwanted attention and contact making the victim fearful for their safety; • sending unwanted gifts; • showing up in places the victim is at; • sneaking into victim's home • following <p><u>Financial abuse</u></p> <ul style="list-style-type: none"> • parasitic lifestyle (financially relying on

<p>themselves (memory or perception) of past or current events/information)</p> <p><u>Vandalism</u></p> <ul style="list-style-type: none"> • damaging victim's property <p><u>Other (specify):</u></p> <hr/> <hr/>	<ul style="list-style-type: none"> • Punch • Hold victim down • Unwanted sexual contact • Rape or forced penetration • Shoot • Stab • Slash • Spit 	<p>vic, drain on their finances)</p> <ul style="list-style-type: none"> • fraudulent financial behaviour that impacts the victim • missing bill payments on accounts in victim's name • stealing from vic or their children
---	--	--

Appendix C

Results of Planned Analyses for the Dynamic Risk Composite

A composite variable was developed by summing the score on all seven dynamic risk factors. The total dynamic risk scale was calculated by summing the ratings of each suspect on the seven dynamic risk factors (ranging from 0 to 2 on the item level, and 0 to 14 on the total scale level; $M = 5.11$, $SD = 2.88$). The internal consistency analysis revealed a Cronbach's α of .51 for the total scale of the dynamic risk factors, indicating poor internal consistency.

A receiver operating characteristic curve (ROC) analysis was used to assess the predictive accuracy of the total dynamic risk score for imminent (< 1 month), short-term (up to 3 months), medium (up to 6 months), and long-term (up to 12 months) time periods. That is, the total dynamic risk score for the index offence was used to predict any recidivism during these time points. Using the full sample, the ROC analysis indicated that the total dynamic risk score had small effect sizes, indicating poor predictive validity for IPV recidivism within 1 month, $AUC = .259$, 95% CI [.4471, .667], 3 months, $AUC = .606$, 95% CI [.519, .693], 6 months, $AUC = .584$, 95% CI [.505, .662], or 12 months, $AUC = .581$, 95% CI [.507, .654].

As the next step, the DeLong et al. (1998) method of comparing AUCs was used to compare the produced AUCs for male and female suspects. The results of these analyses demonstrated no significant gender differences. This method was also used to compare the total dynamic risk score to the researcher-scored ODARA at each time point, finding they are comparable at predicting recidivism at 3 months, but the ODARA

outperforms the dynamic risk total score at every other time point (see Appendix D for results of AUC comparisons).

The ROC analysis can also be used to determine the cut points for offender classification (e.g., low, medium, or high risk for reoffending) on the dynamic risk composite variable. Upon examination of the cut points of the ROC curve to determine the dynamic risk composite score that provided optimal prediction, it was noted that both sensitivity and specificity were maximized at a score of 5.5 out of 14, which was consistent across all four time periods. This cut point was used to create a dichotomous variable for which scores from 0 to 5.4 were coded as “lower dynamic risk” and scores equal to or greater than 5.5 were coded as “higher dynamic risk.” This dichotomous variable was then used to examine the dynamic risk composite’s ability to estimate imminent/short-term risk for IPV recidivism in comparison to the ODARA.

A cox regression survival analysis was used to model the length of time in days between the index IPV event and the first recidivistic event. The covariates were the researcher-scored ODARA classification (entered in Block 1), and the dynamic risk classification (entered in Block 2). A priori power analysis indicated that a sample of 88 was required to detect a significant result with a medium effect (.15), with a power of .90 and an alpha of .05, which was surpassed with an $N = 298$ for this analysis. The ODARA risk classification, in Block 1, significantly contributed to the prediction of time to IPV recidivism, $\chi^2(1) = 18.89, p < .001$. In Block 2, although the overall model still significantly predicted time to recidivism, $\chi^2(2) = 18.56, p < .001$, the dynamic risk composite did not demonstrate incremental predictive improvements beyond the ODARA, ($\beta = .02, p = .69, \text{Exp}(B1) = 1.02$), $\chi^2(1, N = 298) = .16, p = .69$.

The final test of the dynamic risk composite involved examining if the tool was able to reliably detect change in risk over time. The RCI analysis used to assess whether an individual's change in their total dynamic risk score was reliable (i.e., not explained by chance or measurement error). In terms of the total dynamic risk score, the results indicated that 66.7% ($n = 58$) of the suspects had a reliable decrease in their total dynamic risk, 9.2% ($n = 8$) had no reliable change, and 24.1% ($n = 21$) had a reliable increase in their total dynamic risk. As a follow-up to the RCI analyses, a paired-samples t-test was used to determine whether there was a statistically significant mean difference between the total dynamic risk scores for suspects at index and at the first recidivism event. The results indicated that the suspects in this sample who recidivated had higher average total dynamic risk at the index ($M = 5.68$, $SD = 3.13$) than at the first recidivism event ($M = 4.44$, $SD = 2.52$), 95% CI [0.478, 2.004], $t(86) = 3.24$, $p = .002$.

Appendix D

Predictive Validity of the Dynamic Risk Composite

Table D.1

Predictive Validity of Total Dynamic Risk Score for Predicting IPV Recidivism as Measured by ROC Analyses

Follow-Up Period	AUC [95% CI]			% CI Overlap
	Full Sample (<i>N</i> = 298)	Females (<i>n</i> = 99)	Males (<i>n</i> = 199)	
1 month	.259 [.471, .667]	.479 [.246, .712]	.563 [.457, .670]	45.71
3 months	.606 [.519, .693]	.554 [.357, .750]	.595 [.498, .692]	49.36
6 months	.584 [.505, .662]	.536 [.376, .697]	.571 [.481, .661]	56.07
12 months	.581 [.507, .654]	.547 [.392, .702]	.559 [.474, .644]	54.84

Note. IPV = Intimate partner violence; ROC = Receiver operating curve; AUC = Area under the curve; CI = Confidence intervals; % CI Overlap = Percent of overlap between confidence intervals for males and females

Table D.2

Predictive Validity of Total Dynamic Risk Score for Predicting IPV Recidivism as Measured by ROC Analyses

Follow-Up Period	AUC [95% CI]		% CI Overlap
	ODARA	ODARA + DRF Score	
1 month	.259 [.471, .667]	.666 [.571, .762]	24.81
3 months	.606 [.519, .693]	.674 [.592, .756]	29.88
6 months	.584 [.505, .662]	.663 [.590, .736]	23.76
12 months	.581 [.507, .654]	.659 [.591, .727]	22.26

Note. IPV = Intimate partner violence; ROC = Receiver operating curve; AUC = Area under the curve; CI = Confidence intervals; DRF = Dynamic Risk Factor; % CI Overlap = Percent of overlap between confidence intervals for ODARA and Total Dynamic Risk Score. *N* = 298.

CURRICULUM VITAE

Candidate's full name:

Erin Jacoba deJong

Universities attended (with dates and degrees obtained):

Carleton University, 2013, Bachelor of Arts (Honours)

University of New Brunswick, *in progress*, Doctor of Philosophy

Publications & Conference Presentations:

Campbell, M. A., **deJong, E.**, McTague, J., Totten, A., & Connell, B (2019, May 20-21). *Factors influencing police use of an intimate partner violence risk appraisal tool and its influence on police response* [Conference presentation]. The American Society for Evidence-Based Policing Conference, Cincinnati, Ohio.
<https://www.americansebp.org/>

Campbell, M. A., **deJong, E.**, Richard, A., Fairholm, J., & Coy, S. (2017, June 7-10). *Impact of bullying prevention: An examination of change in participants of the Canadian Red Cross Beyond the Hurt program* [Conference presentation]. The Canadian Psychological Association National Convention, Toronto, Ontario.

deJong, E. (2013, April 6). *Sex in the courtroom: A jury decision-making study* [Poster presentation]. The Psychology Undergraduate Research Event, Ottawa, Ontario.

deJong, E. (2015, June 2-4). *Police Response to intimate partner violence in New Brunswick* [Conference presentation]. The 3rd North American Correctional and Criminal Justice Psychology Conference, Ottawa, Ontario.

- deJong, E.** (2016, June 10). *Beyond the Hurt outcome evaluation: What works and future directions* [Webinar presentation]. The Beyond the Hurt Atlantic Speaker's Series in association with the Canadian Red Cross and the Medavie Health Foundation.
- deJong, E., Campbell, M. A., & McTague, J.** (2019, May 31 – June 2). *The influence of ODARA risk information on police response to IPV calls* [Conference presentation]. The 4th North American Correctional & Criminal Justice Psychology Conference, Halifax, Nova Scotia. <https://cpa.ca/naccjpc/>
- deJong, E., Campbell, M. A., & Totten, A.** (2018, March 22-23). *From turning a blind eye to evidenced-based policing: The past, present, and future of criminal justice responding to IPV* [Conference presentation]. The Canadian Domestic Violence Conference 5, Halifax, NS.
- deJong, E., Huang, A., Barrington, J., Hale, M., & Whelan, B.** (2021, June 10). *Virtual trials: Clinician's perspectives on facilitating online skills-based groups* [Conference presentation]. The Canadian Psychological Association's 82nd Annual National Convention, virtual conference. <https://convention.cpa.ca/>
- deJong, E., McTague, J., & Campbell, M. A.** (2020, March 7-10). *An evaluation of frontline police use of an intimate partner violence risk assessment tool* [Conference presentation]. The American Psychology-Law Society Annual Conference, New Orleans, Louisiana. <https://ap-ls.org/2020-conference>
- deJong, E., Richard, A., Coy, S., & Campbell, M. A.** (2016, May 6). *The impact of bullying prevention: An examination of change in participants of the Canadian Red Cross Beyond the Hurt program* [Conference presentation and panel

discussion]. The Atlantic Criminal Justice and Professional Practice Conference, Saint John, New Brunswick.

deJong, E., Totten, A., & Campbell., M. A. (2018, March 22). *A history of police response to intimate partner violence and present day application of evidence-based policing* [Conference presentation]. The 5th Annual Canadian Domestic Violence Conference, Halifax, Nova Scotia.

McManus, L., **deJong, E., & Maeder, E.** (2014, March 15). *A test of gender-crime congruence on juror decision-making* [Conference presentation]. The American Psychology-Law Society Conference, New Orleans, Louisiana.

Academic Awards:

- 2018 Department of Psychology Discretionary Award | UNB
- 2017 Faculty of Arts Ph.D. Graduate Assistantship | UNB
- 2017 Student Travel Award | CPA Student Section
- 2016 Snodgrass Research Proposal Award | UNB
- 2016 Faculty of Arts Ph.D. Graduate Assistantship | UNB
- 2015 School of Graduate Studies Grant for Conference Presentations | UNB
- 2015 Faculty of Arts Ph.D. Graduate Assistantship | UNB
- 2014 Faculty of Arts Ph.D. Graduate Assistantship | UNB
- 2013 Senate Medal for Outstanding Academic Achievement | Carleton University
- 2013 Faculty of Arts M.A. Graduate Assistantship | UNB Faculty of Arts
- 2013 Highest Honours Distinction in the Faculty of Arts | Carleton University
- 2012 Dean's Honour List | Carleton University
- 2012 Elinor Burwell Scholarship in Psychology | Carleton University

- 2011 Dean's Honour List | Carleton University
- 2011 Harry S. Southam Scholarship | Carleton University
- 2010 Dean's Honour List | Carleton University
- 2010 Chalmers Jack MacKenzie Scholarship | Carleton University
- 2009 President's Scholarship | Carleton University
- 2009 Johnson Scholarship | Carleton University
- 2009 Ontario Scholar Award | Ontario Ministry of Education