

————— Research Report —————

**Family Violence Programming:
Treatment Outcome for Canadian
Federally Sentenced Offenders**

Ce rapport est également disponible en français. This report is also available in French. Pour obtenir des exemplaires supplémentaires, veuillez vous adresser à la Direction de la recherche, Service correctionnel du Canada, 340, avenue Laurier ouest, Ottawa (Ontario) K1A 0P9. Should additional copies be required, they can be obtained from the Research Branch, Correctional Service of Canada, 340 Laurier Ave., West, Ottawa, Ontario, K1A 0P9.

2008 N° R-174

**Family Violence Programming:
Treatment Outcome for Canadian Federally Sentenced Offenders**

Lynn Stewart¹, Natalie Gabora¹, Randy Kropp²

&

Zina Lee²

¹Correctional Operations and Programs, Reintegration Programs Division
Correctional Service Canada

²B.C. Institute Against Family Violence

October 2005

ACKNOWLEDGEMENTS

The outcome study portion of this program evaluation was completed under contract by Dr. Randy Kropp and Zina Lee from the B.C. Institute Against Family Violence. We would like to thank the following research assistants for their role in preparing the data set: Dayna Medveduke, Kate Rossiter, Corrine MacPherson and Neala Kelly. In addition, we wish to acknowledge the work of the skilled facilitators who deliver the programs and work with the offenders in working toward their goal of developing healthy, non-violent relationships.

This study was funded in part by the Departments of Health and Justice, and by the Correctional Service of Canada.

EXECUTIVE SUMMARY

Since 2001, CSC has implemented programs for male offenders who have a history of violence against their female partners. These programs were designed to conform to the best practices in the effective corrections literature. The program menu includes a moderate-intensity program for offenders with less extensive histories of partner abuse, a high-intensity program for higher risk offenders, a maintenance program for participants who have completed either program, and a treatment primer designed to enhance motivation of potential participants of the program. These programs were accredited by an international accreditation panel in 2001. This paper describes the programs and the results of a preliminary evaluation examining their effectiveness.

The evaluation confirms that program participation significantly reduces attitudes that support violence against women and increases pro-social skills related to non-abusive relationships. Treatment effects, as measured by pre- and post-treatment change scores on measures assessing the significant targets of the program, were high to moderate. Feedback solicited from parole officers supervising graduates generally reported positive changes in behaviour and attitude associated with treatment. Participants themselves overwhelmingly stated that they found the program useful and reported that they will be able to use the skills to reduce their risk of partner violence.

A six month post-release follow-up indicated that program completion in the high and moderate intensity family violence prevention programs reduced spousal violence by 69% relative to the non-treatment comparison group. There was a 47% decrease in any violent crime. The programs did not have a significant impact on non-violent crime.

The paper provides a profile of offenders who dropped out of treatment as well as those who were rated at post-program as having achieved an overall positive program performance. Of note, ethnic group membership, risk level and borderline personality diagnosis were not related to overall program performance. Only two factors (Antisocial Personality Disorder and a rating on the extent of offenders' abusive attitudes towards women) were weakly related to spousal assault recidivism, but this effect may be attenuated by the small number of program graduates who could be followed and the low base rate of reoffending. Overall, several disparate indicators converge to point to these programs as promising interventions to reduce abusive attitudes and intimate violence among male correctional populations.

TABLE OF CONTENTS

| | |
|--|-----|
| ACKNOWLEDGEMENTS | i |
| EXECUTIVE SUMMARY | ii |
| TABLE OF CONTENTS..... | iii |
| LIST OF TABLES | iv |
| INTRODOCUTION | 1 |
| FAMILY VIOLENCE PROGRAMS | 3 |
| Programs to Address Family Violence Within the Correctional Service of Canada..... | 3 |
| CURRENT STUDY..... | 6 |
| METHODOLOGY | 7 |
| Participants | 7 |
| Measures..... | 9 |
| Post-Release Follow-Up..... | 12 |
| RESULTS | 14 |
| Measures..... | 14 |
| Quizzes | 17 |
| Parole Officer Feedback | 18 |
| Participant Feedback..... | 19 |
| Recidivism..... | 20 |
| Factors Related to Success in Treatment..... | 22 |
| DISCUSSION | 27 |
| REFERENCES | 30 |

LIST OF TABLES

Table 1. Moderate- and High-Intensity Program Effect Sizes..... 15

Table 2. Parole Officers’ Judgments on Goal Attainment Scale Items – Post Treatment 19

Table 3. Recidivism Rates for Treated Versus Untreated Groups..... 20

Table 4. Recidivism Odds Ratios for Treated Versus Untreated Groups 21

Table 5. Overall Effect (χ^2) and Relative Effects (*OR*) of Treatment Completion and Intensity on
Recidivism..... 22

INTRODUCTION¹

Outcome studies evaluating the impact of treatment programs for male batterers have provided equivocal results. The conclusions drawn from review articles (e.g, Bennett and Williams, 2004) point out that the programs appear to be less effective when more rigorous experimental designs are applied (i.e., Dunford, 2000; Feder & Forde, 2000; Taylor, Davis, & Maxwell, 2001), although experimental designs are also not without their critics (see Gondolf, 2001a; Dobash & Dobash, 2000). Quasi-experimental designs where the treatment group is compared to a non-equivalent comparison group have generally reported small positive treatment effects (Babcock, Green & Robie, 2004)). Studies that evaluate the impact of the program based on participants' level of violence prior to and after treatment have also largely concluded that levels of violence diminish after treatment, but the treatment effect is difficult to disentangle from the deterrent impact of arrest and supervision. A recent meta-analytic study summarising the impact of several evaluations that met a standard of research design concluded that treatment participation produced, at best, a weak effect (in the order of .10 which means a reduction of 5% in recidivism) (Babcock et al., 2004).

Gondolf, however, has stated that the conclusions drawn from evaluations of domestic violence programs may be unduly pessimistic because they do not examine the full impact that program participation may have. He and his team have looked at a number of convergent outcome factors, in addition to recidivism, that reflect treatment effect. For example, in their multi-site evaluation of three treatment programs, they report that nearly 67% of the partners of treated men at the 15-month, 30-month and 48-month follow-ups indicated that they were "better off." At the 30-month and 48-month follow-ups, nearly 85% of the women reported feeling "very safe" and "very unlikely to be hit again." The women also reported a change in the men's behaviours and attitudes which they attributed to the programs. The participants themselves reported that the methods they use to avoid being abusive are primarily techniques learned in the programs (Gondolf, 2001).

Several shortcomings have been identified in the design of domestic violence programs and their evaluations. Failure to identify or measure impact of treatment on intermediate treatment objectives, failure to identify a viable comparison group, failure to ensure program

¹ Findings from this report were presented at the International Conference Against Family Violence, Banff, Alberta, Canada, October 2005.

integrity and high attrition rates are some of the problems in program delivery and program evaluation.

High dropout rates ranging up to 80% of those who originally present for treatment make program evaluation difficult because the participant characteristics related to drop out are often also related to characteristics that affect program outcome. Typically, drop outs from correctional programs have much poorer outcomes than program completers (Dowden & Serin, 2001). Most of the published outcome literature fails to report on the progress of participants on the targets of the program, leaving unanswered the question of whether the offenders have made gains on objectives theoretically related to domestic violence and whether the extent to which offenders demonstrate specific participant abilities or skills contributes to outcome. Flawed evaluation designs have often made it difficult to draw conclusions; few studies use a random assignment design and most have no comparison group or have not adequately established the equivalence of the comparison group. Finally, many service agencies delivering programs do not have systems in place to ensure a certain level of staff training and the ongoing integrity of program delivery.

This paper will present the preliminary outcome evaluation of two family violence programs delivered nationally to federal offenders in the Correctional Service Canada. The program curriculum, assessment procedures and program implementation strategy were designed to meet international standards of recognised in the “effective corrections” literature (see Andrews & Bonta, 1998; Gendreau, 1996; Gendreau & Andrews, 1996).

FAMILY VIOLENCE PROGRAMS

Programs to Address Family Violence within the Correctional Service of Canada (CSC)

The risk factors associated with intimate partner violence are similar to those associated with general criminal behaviour (Hanson & Wallace-Capretta, 2000; Hart et al., 1994). It is not surprising, therefore, that an estimated 40% of male federal offenders have some history of violence against their intimate female partners.

The Correctional Service of Canada (CSC) has mandated treatment for all male offenders who have been identified as being at continued risk to be abusive in intimate relationships. An initial assessment at reception screens all offenders for their risk for future violence in their intimate relationships. Based on their risk level and the extent of the violence in intimate relationships, they will be referred to either a moderate- or a high-intensity family violence prevention program.²

CSC has endorsed a process of program accreditation that sets out criteria that correctional programs must meet in their design and implementation. These criteria were identified through a review of the effective corrections literature and were officially endorsed by CSC's Commissioner. The eight criteria that must be demonstrated are: the program must be based on an explicit, empirically-based model of change, target criminogenic need (i.e., needs empirically linked to criminal behaviour), use effective methods, include skills development; address responsivity issues; provide sufficient intensity of service; provide a continuity of care, and ensure ongoing monitoring and evaluation. The accreditation process is overseen by an international panel of experts in corrections that makes its recommendation to accredit to the commissioner. The moderate- and high-intensity family violence prevention programs were accredited by the international panel in 2001. Since then, they have both been implemented across the country.

The moderate-intensity program is a 25-session program for moderate risk offenders, delivered both in the community and in the institutions. High-risk offenders attend the 78-session high-intensity program which is delivered only in the prisons.

Both programs adopt the nested ecological model (Dutton, 1995) that explains intimate

² Offenders assessed as moderate risk on the Spousal Assault Risk Assessment (SARA) Guide with one incident of abuse were referred to the moderate program; offenders assessed as high risk on the SARA with more than one incident of abuse were referred to the high-intensity program.

violence as a multi-determined behaviour. The nested model, which is derived from the work of developmental psychologists and ethologists, considers the interactions between four levels of influence on the target behaviour; in this case, violence and abuse against women. At the most general level, elements of the broad social context (Macrosystem) provide tacit or explicit support for men's violence in general and violence toward women in particular; at the next level (Exosystem), the behaviour is also influenced by the presence or absence of social structures such as prosocial supportive, or alternatively, peers and family members who endorse the abuse of women. The Microsystem level considers the immediate context for the violence and abuse, such as the amount of conflict between the couple, their communication pattern and each spouse's method of coping with conflict. Finally, at the Ontogenic level, the model considers the impact of the perpetrator's intrapsychic features such as his history of being a victim or a witness to violence, the pattern of attachment to early caregivers, and individual differences related to problems in self-regulation or, in some cases, even neurological impairment.

The CSC programs attempt to address factors at all levels defined by the model. The programs' initial modules establish the offenders' motivation for change and educate offenders on the range of abusive behaviours and factors that contribute to their abusive patterns towards women. Other modules train offenders on cognitive restructuring of attitudes and beliefs that condone the abuse of women and in skills that address the management of emotions of jealousy, anger and fear of relationship loss that are associated with abuse. Participants are trained on key social and communication skills that underlie healthy relationships such as negotiation and responding appropriately to criticism. Later modules involve the offenders in the development of relapse prevention plans that include the planning of coping strategies to avoid or manage high-risk situations, the identification of people to avoid who contribute to the risk of further abuse and, conversely, the identification of a support network who will assist in maintaining a commitment to healthy relationships. The high-intensity program includes a more detailed examination of the influences on the development of abusive patterns through autobiographies and a longer module on parenting.

Both programs incorporate a stage-of-change approach that recognises that not all offenders will be equally ready to address their history of abuse against their partners. A stage-appropriate treatment primer is available for offenders who meet the referral criteria, but are not ready to attend the program. As well, the introductory section of the program itself gives the

offenders time to evaluate their goals for change and assess how they want to change.

The curriculum is delivered by trained facilitators whose adherence to the manual and to the principles of effective correctional program delivery is monitored through videotape review of key sessions. Detailed quality review reports of their work are conducted by regional trainers who provide feedback on their program delivery to assist in professional development and also insure that they meet a specified standard in their assessments and final program reports.

Programs are delivered by teams made up of a male and a female co-facilitator to a group of up to 12 offenders. Sessions are 2-3 hours, delivered 3 to 5 times per week. At least 3 (moderate-intensity program) to 10 (high-intensity program) individual sessions are conducted with one of the facilitators assigned as a primary counsellor. Offenders participate in a detailed assessment of their pre-program attitudes and skill levels and are reassessed upon completion of the program.

CURRENT STUDY

The following paper reports on an evaluation of the impact of both the high- and moderate-intensity programs including an assessment of the post-program change in attitude and skill level and the result of a limited follow-up of program participants who had been released from prison. The study also points to what factors are most significantly related to key program outcome indicators: program completion, overall development of skills and prosocial attitudes, and avoidance of intimate partner violence upon release into the community.

METHODOLOGY

Participants

Participants were 572 male offenders incarcerated at various federal institutions across Canada who participated in either the high- or moderate-intensity family violence prevention programs between 1999 and 2003. Approximately 41% ($n = 237$) of the offenders had participated in the high-intensity program and 59% ($n = 335$) had participated in the moderate-intensity program. The mean sentence length of the total treatment group was 4.24 years ($SD = 2.99$) and ranged from 2 to 22 years. These sentences were not necessarily for a spousal assault; but, to meet the referral criteria, the participants had to have at least one incident of assault of an intimate partner in their personal history. This abusive behaviour need not have resulted in a conviction. There was no significant difference between the high- and moderate-intensity participants with respect to sentence length or on any of the demographic variables. Overall, the men ranged in age from 20 to 69 years, with an average age of 37 years ($SD = 9.50$). Most of the offenders were Caucasian (62%), with the remainder Aboriginal (25%) or of "Other" ethnicity (13%) which includes a variety of ethnic groups. Only 25% of the offenders had a high school education. According to the Shipley Institute of Living Scale (Zachary, 1991), the mean estimated IQ of the sample was 96.91 ($SD = 20.59$). Forty-seven percent of participants had significant problems with alcohol abuse as assessed on the Computerised Lifestyle Assessment Instrument (CLAI). Of note, 79% of the high-intensity and 60% of the moderate program participants reported being under the influence of drugs or alcohol when they assaulted their partners.

The high-intensity program participants were assessed for personality disorder in order to determine whether there was a differential response to treatment for the Borderline or Antisocial Personality offenders. This assessment was not completed on the moderate-intensity program participants. The mean score on the measure of Borderline Personality Disorder (BPO) was 62.04 ($SD = 18.36$). In comparison, the mean BPO score for the sample of wife assaulters from Dutton and Haring's study (1995) was 71.3 ($SD = 17.1$). Oldham et al's (1985) mean score for diagnosed borderlines was 74.8 (Dutton & Haring, 1995). The CSC scores were actually closer to the mean scores of Oldham et al's non-borderline sample and Dutton's control group (61.3 and 60.0, respectively). Using mean scores reported by Oldham et al. (1985), we estimate that one quarter (26%) of the CSC sample met the diagnosis of BPO.

Missing data on some of the criterion items made it more difficult to determine the percentage of offenders who met the criteria for Antisocial Personality Disorder (APD). However, given the data we have, we can estimate that about 75% of the population would be diagnosed with APD.

Not surprisingly, given that the offenders were federal offenders, they had substantial histories of assaultive behaviour towards both their partners and others. Most (57%) reported having been in 10 or more fights; most of these were with males (73%). About 14% reported having been in more than 50 fights.

Although participants in both programs admitted to substantial physical and emotional abuse of their partners, the high-intensity program participants acknowledged more, and more serious, abuse, than did the moderate intensity participants, confirming that the offenders were allocated to the appropriate treatment option. For example, the high-intensity program participants were more likely to admit to hitting (72% as opposed to 53%), choking (31% as opposed to 19%) and use of a weapon (40% as opposed to 19%) than the moderate-intensity participants. Thirty percent of the high and 17% of the moderate-intensity participants admitted having assaulted their partners more than 5 times. Both treatment groups also admitted to equal levels of emotional abuse ranging from yelling (89-90%), swearing (86-88%) and threatening (58-59%) to isolating her (17-25%) and property damage (47-49%). Participants in both programs were less likely to admit to sexual abuse, with 16% of the high- and moderate-intensity participants admitting to forcing sex or pressuring sex when their partner did not want it. Their partners had called the police because of their abuse in two-thirds (67%) of the cases for the high-intensity and 61% of the moderate-intensity participants; in 22% of the high-intensity cases and 10% of the moderate-intensity, the police were called 4 or more times. A majority of the men in the high-intensity program (58%) and 41% of the moderate-intensity participants had been assaultive towards more than one partner. As expected, the offenders' social histories included experience of personal victimisation and family disruption: most of the high-intensity participants (61%) and half of the moderate-intensity participants had witnessed the abuse of their mothers and two thirds of the men from both programs had themselves been abused as a child. More than half of the participants in both programs reported that family members had criminal records.

On a more positive note, almost all of the offenders in both the moderate and high

program at the initial interview (96-99%) reported believing that the program could be “somewhat” or “very useful” for them. This is reflected in the completion rates for the program. Relative to other batterer intervention programs reported in the literature (Daly et al., 2001), the dropout rates were respectable: only 18% of offenders dropped out of the high-intensity program and 14% of offenders dropped out of the moderate-intensity program.

Measures

In addition to a number of profiling measures that assessed the offenders’ level of psychopathology, all offenders completed a battery designed to assess the impact of the program on the program targets: attitudes towards women abuse, and lack of skill in managing emotions and social situations and coping with high-risk situations. To provide divergent sources of evaluation, we chose a variety of measures including self report, scenario-based skills assessment (vignettes) and facilitator ratings. Correlational analyses confirmed that one of the strengths of including the vignettes and rating scales in the battery was that, unlike the self report measures, the scores were unrelated to offenders’ deliberate attempts to “fake good” as measured by the Paulhus Deception Scale.

The pre- and post- measures are described briefly below. Norms for the various measures can be obtained by writing to the first author of this report.

Interpersonal Relationship Scale. The Interpersonal Relationship Scale (Hupka & Rusch, 2001) is a 27-item self-report measure that assesses six aspects of jealousy: threat to exclusive companionship, self-deprecation/envy, dependency, sexual possessiveness, competition and vindictiveness, and distrust. Each item is rated on a 6-point scale (1 = *strongly agree*, 6 = *strongly disagree*), with higher scores indicating lower levels of jealousy.

Abusive Relationships Inventory. The Abusive Relationships Inventory (Boer, Kroner, Wong, & Cadsky, undated) is a 33-item self-report measure that assesses 4 scales related to abusive relationships: rationale for hitting, need for control, legal entitlement, and batterers’ myths. Each item is rated on a 7-point scale (1 = *strongly disagree*, 7 = *strongly agree*), with higher scores indicating more negative attitudes toward relationships.

Relationship Style Questionnaire. The Relationship Style Questionnaire (Griffin & Bartholomew, 1994) is a 30-item self-report measure designed to assess the four attachment styles: secure, fearful, preoccupied, and dismissing. Each item is rated on a 5-point scale (1 = *not at all like me*, 5 = *very much like me*).

Personal Reaction Questionnaire. The Personal Reaction Questionnaire (Blackburn &

Fawcett, 1999a; Blackburn & Fawcett, 1999b) is a 125-item, multi-trait, self-report inventory designed to measure cognitive, affective, and behavioural dispositions of relevance to offender populations. It provides a brief assessment of deviant personality traits and is intended to facilitate the identification of an offender's problems and the planning and monitoring of clinical intervention. Each item has a dichotomous response choice (yes/no), and the measure assesses 8 personality traits: resentment, aggression, self-esteem, avoidance, paranoid suspicion, extraversion, deviance, and self-control.

Paulhus Deception Scale (PDS). The Paulhus Deception Scale (Paulhus, 1990) is a 40-item self-report measure designed to assess two subscales of socially desirable responding: self-deceptive enhancement and impression management. Each item is rated on a 5-point scale (1 = *not true*, 5 = *very true*). Self-deceptive enhancement refers to a person's tendency to give honest but exaggerated positive self-reports; whereas impression management refers to a person's tendency toward purposeful manipulation of answers to appear more socially acceptable.

Borderline Personality Organization. The Borderline Personality Organization instrument (Oldham, Clarkin, Applebaum, Carr, Kernberg, Lottermen, & Haas, 1985) is a 30-item self-report measure that assesses components of a borderline personality. Each item is rated on a 5-point scale (1 = *never true*, 5 = *always true*), and items are summed to yield three subscales: loss of reality, primitive defences, and identity confusion.

Relapse Prevention Test, Family Violence Vignettes and Empathy Scales. These measures were developed within CSC to assess skills and attitudes targeted by the program (CSC, 2001). They are based on structured interviews in which the respondent is asked how he would respond in a number of situations related to family violence. There are two versions of the tests (Version A and Version B), with one version administered at pre-treatment and the second at post-treatment. Administration of versions A and B is alternated across offenders to counterbalance the learning effect and level of difficulty. More detailed descriptions of these measures follows.

Relapse Prevention Test. There are a total of eight scenarios, two for each of four potential risk situations. Four of the scenarios are administered at pre-treatment and four are administered at post-treatment. The test is designed to assess an offender's: (1) recognition of situations leading to violence; (2) effective and systematic use of skills such as problem solving, emotions management, and network support; and (3) ability to evaluate the effectiveness of his solutions. In addition, it requires the facilitator to count how many potentially effective responses the

offender suggested. Higher scores indicate that the respondent has good recognition and application of skills.

Family Violence Vignettes. The Family Violence Vignettes examine responses to aspects of relationship violence in five situations: jealousy, employment and sexual issues. Higher scores indicate that the respondent assesses his partner with respect and as an equal, and uses appropriate behaviour.

Empathy Scales. The Empathy Scales involve a structured interview in which the participants are given a number of situations and asked how they would respond. The situations represent a partner in distress, a child in distress, and a person outside the family in distress. Responses are scored on a 3-point scale to assess perspective taking, sincerity of affect, and coping with distress. Total scores range from 0 – 72, with higher scores indicating better empathy skills.

Treatment Readiness, Treatment Responsivity, and Treatment Participation and Gain Scales. The Treatment Readiness, Treatment Responsivity, and Treatment Participation and Gain Scales (Serin & Kennedy, 1997) are structured interviews designed to assess treatment readiness and responsivity factors that could affect an offender's response to treatment.

Goal Attainment Scale-FV. The Goal Attainment Scale-FV is a rating scale of the offender's behaviour and attitudes on 14 treatment goals associated with the CSC family violence prevention programs. Each goal is rated on a 5-point scale (-2 = "very risky" attitude or very low skill achievement; 0 = minimal acceptable attitude or skill achievement; +2 = very prosocial attitude or high skill achievement). The scale is completed for each participant with input from both facilitators to ensure greater reliability. The scale was developed within CSC, based on a scale for sex offender program participants (Hogue, 1993).

University of Rhode Island Change Assessment -Batterers. (URICA-B, Levesque, Gelles, & Velicer, 2000). This instrument, based on the more generic URICA, measures offender's stage of change with respect to their willingness to stop abusive and violent behaviour against an intimate partner. The self report inventory presents statements characteristic of various stages of change (precontemplation, contemplation, etc) and the offender rates the extent to which the item is like him on a 5 point scale. The measure is administered during the initial interview for both the moderate and high intensity offenders.

Following the initial interviews and assessment offenders were rated by facilitators on a 5-point scale on motivation for the program, extent of abusive thinking, and strategies for managing

emotion and resolving conflict. The results of this assessment were provided to the offender so that he could work with the facilitator in agreeing to the goals of treatment. During the program, participants also completed quizzes at the end of every module to determine the extent to which they could understand and apply the material in the program.

Feedback was solicited from the participants and parole officers supervising the program participants on their opinion of the impact of the program. From the offenders we requested detailed feedback on specific aspects of the program and overall feedback on the general utility of the program. The feedback from the offenders was collected using confidential questionnaires completed anonymously at the end of the program.

Parole officers who were supervising program participants or graduates, either in the community or institutions, were mailed a letter by independent contractors, explaining the purpose of the study and requested to participate in a short interview. The interview questions asked for their observations of the impact of the family violence prevention programs on offenders' attitudes and behaviour since their completion of treatment (note that we did not interview parole officers about offenders who had not participated in treatment). We contacted a random group of parole officers that represented 65% of offenders who participated in either the high or moderate intensity family violence program. Of these parole officers, only 19% (N = 74) agreed to participate in a confidential interview. The interview was structured to elicit ratings of the offenders on key factors that related to abusive intimate behaviour and risk for further abusive behaviour.

Post-Release Follow-Up

The final phase of the evaluation was a post-release follow-up of 246 family violence offenders falling into 2 main groups:

1. *Treated* offenders who had completed either the high-or moderate- intensity family violence program and had been released. This group included 84 participants from the high-intensity program and 76 participants from the moderate intensity program.
2. *Untreated* offenders who met the admission criteria for the programs (i.e., they had a history of spousal violence and a SARA rating of moderate or high) but who, for various reasons, did not begin or complete the program. This group was comprised of 35 offenders who were eligible for the high intensity program: 13 who received no treatment for administrative reasons (e.g., institutional transfer) and 22 who had dropped out of treatment. The group also included 51 offenders who were eligible for the moderate intensity program (22- no treatment,

29 drop outs).³ (Note: We had insufficient background information to control for potential differences between the groups on variables such as sentence length, marital status, history of violence and so forth. However, for these analyses we used all available offenders who were released, so it is unlikely that there were any systematic differences between groups).

To be eligible for this phase of the evaluation, offenders were also required to have spent at least 6 months in the community following their release. Treated offenders in this sample spent on average 369 days at risk in the community, whereas untreated offenders averaged 462 days at risk. The difference between the two groups was not statistically significant. Despite our best efforts, we were unable to contact enough spouses of the released program completers, so we had to rely on Canadian Police Information Centre (CPIC) and CSC Offender Management System (OMS) records to record new criminal offences and supervision violations following release.

Initially new offences were tallied in five categories: (1) confirmed spousal violence, including any actual, attempted or threatened physical or sexual violence against a spouse; (2) other spousal-related crime including harassing behaviours and breaches of no-contact conditions; (3) non-spousal violence; (4) non-violent crime; and (5) non-spousal-related conditional release violations. Due to the low base rate of re-offending, however, we chose to create three nested offence categories for our analyses of recidivism. They are: *Spousal Violence*, which includes all spousal-related crime; *Any Violence*, which includes spousal violence and other violence; and *Any Infraction*, which includes all violent and non-violent crime, regardless of whether or not the crime resulted in arrest.

³ Preliminary analyses indicated that there were no differences in recidivism rates between those who had received no treatment and those who had dropped out of treatment. Therefore, we chose to combine these offenders into the group labelled “Untreated Offenders” in order to increase the sample size for subsequent analyses. The reasons for dropping out of treatment varied: of the 51 who dropped out, about 30% dropped out because they were transferred to another institution, had a conflict with scheduling employment, or had problems understanding the language of instruction, variables that are not theoretically linked to risk to reoffend).

RESULTS

Measures

We compared the pre-treatment and post-treatment performance of the program participants on the test measures by conducting a series of t-tests, applying a Bonferroni correction for the number of analyses. Almost every comparison found that offenders improved significantly after participation in treatment. In general, upon treatment completion, offenders in both the high- and moderate-intensity programs evidenced significantly lower levels of jealousy (Interpersonal Relationship Scale), fewer negative attitudes about relationships (Abusive Relationships Inventory), better recognition and application of relapse prevention skills (Relapse Prevention Test), more respect for their partners (Family Violence Vignettes), greater treatment readiness and responsivity (Treatment Readiness Scales), and more engagement in positive behaviours and attitudes (Goal Attainment Scale).

Given the number of comparisons, Table 1 presents only the calculated effect sizes (Cohen's *d*) for the comparisons instead of presenting the average pre- and post-treatment scores for all of these measures. This allowed us to standardize the change scores so that they could be compared directly. The effect size is an estimate of the level of change from pre to post treatment. Cohen's *d* may be interpreted as follows: an effect size of .3 is considered "small"; an effect size of .5 is considered "moderate"; and an effect size of .8 is considered "large." Effect sizes exceeding 1.0 are considered exceptionally high. According to these guidelines, most of the measures had medium to large effect sizes, with the exception of the Interpersonal Relationship Scale and the Abusive Relationships Inventory, which had small effect sizes. Offenders were tested only once on the Relationship Style Questionnaire, the Personal Reaction Questionnaire, and the Borderline Personality Organisation measures to profile participants and determine their relative response to treatment. The results for these measures will be presented as they relate to key outcome indicators. The Paulhus Deception Scale was used to determine the extent to which the outcome on other measures is correlated with its scales on impression management. For the complete outcome statistics on these measures please refer to Kropp and Lee (2004). The Treatment Readiness Scale was used for the high-intensity participants only.

Table 1. Moderate- and High-Intensity Program Effect Sizes

| Scales | Effect Size Moderate Program (<i>n</i> = 241 – 282)⁴ | Effect Size High Program (<i>n</i> = 110 – 156)⁵ |
|--|--|--|
| <i>Interpersonal Relationship Scale</i> | | |
| Threat to Exclusive Companionship | .43** ⁶ | .59** |
| Self-Deprecation/Envy | .33* ⁷ | .48** |
| Dependency | .52** | .66** |
| Sexual Possessiveness | .30* | .27* |
| Competition and Vindictiveness | .56** | .78** |
| Distrust | .13 | .25* |
| <i>Abusive Relationships Inventory</i> | | |
| Rationales for Hitting | .37** | .36** |
| Need for Control | .56** | .67** |
| Legal Entitlement | .38** | .32** |
| Batterers' Myths | .37** | .52** |
| <i>Relapse Prevention Test</i> | | |
| Question 1 | .84** | 1.05** |
| Question 2 | .83** | 1.23** |
| Question 3 | .80** | .70** |
| Question 4 | .79** | 1.18** |
| Total Score | 1.09** | 1.22** |
| <i>Family Violence Vignettes</i> | | |
| Jealousy | .57** | .18 |
| Employment and Finances | .75** | .53** |

⁴ The size of the group varies because of missing data.

⁵ The size of the group varies because of missing data.

⁶ Probability of the effect due to chance is less than 1 time out of 1000 ($p < .001$)

⁷ Probability of the effect due to chance is less than 1 time out of 100 ($p < .01$)

| | | |
|--|--------|--------|
| Rejection | .53** | .53** |
| Control Issues | .60** | .67** |
| Sexual Issues | .66** | .36** |
| Interpretation | .73** | .65** |
| Behavioural Response | .86** | .83** |
| Total Score | .88** | .71** |
| <i>Empathy Scales</i> | | |
| Perspective Taking | .58** | .69** |
| Affect | .47** | .65** |
| Coping with Distress | .63** | .94** |
| Partner Centred Scenarios | .52** | .71** |
| Child Centred Scenarios | .38** | .73** |
| Person Outside Family | .75** | .66** |
| Total Score | .60** | .84** |
| <i>Goal Attainment Scale (Pre- vs. Post-Treatment)</i> | | |
| Acceptance of Responsibility for Abusive/ Violent Behaviour | 1.01** | 1.18** |
| Acknowledges Use of Power and Control Tactics | 1.20** | 1.32** |
| Shows Empathy for Victims | .96** | .92** |
| Extent of Skills Development | .99** | 1.27** |
| Minimizes Consequences | .91** | 1.16** |
| Understands Life Style Dynamics | 1.07** | 1.20** |
| Understands Abusiveness Pattern | 1.30** | 1.38** |
| Identify Relapse Prevention Concepts | 1.21** | 1.54** |
| Discloses Personal Information | .92** | .94** |

| | | |
|---|--------|--------|
| Participation in Treatment | .80** | .73** |
| Motivation to Change Behaviour | .79** | .65** |
| Overall Quality of Plan on Release | 1.12** | 1.10** |
| Overall Participation and Program Performance | .83** | .85** |
| <i>Treatment Readiness Scales (Pre vs Post-treatment for high participants only)</i> | | |
| Treatment Readiness | | .78** |
| Treatment Responsivity | | .77** |

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

Quizzes

The content quizzes were designed to determine whether offenders understood the key concepts and could apply them to their situation. The scores for both the moderate- and high-intensity programs indicated that most offenders understood most of the content quite well. The mean average scores on the quizzes were between 65% and 77% for the high-intensity program and between 63% and 77% for the moderate-intensity program.

Several ethnicity differences emerged in the results of the quizzes. In the high-intensity program, Caucasians performed better than Aboriginals on quizzes on modules 2, 6, 8, and 9; and both Caucasians and Aboriginals performed better than offenders of “Other” ethnicity on the quiz on module 1. On the moderate-intensity module quizzes there were no significant differences between any of the groups, with the exception of module 2 in which Caucasians scored significantly better than both Aboriginals and offenders of other ethnicity.

A series of tests (Analysis of Variance; ANOVA) were conducted to examine whether there were any further differences between three ethnicity groups (Caucasian, Aboriginal, and Other) on various factors and on outcome. Post-hoc (Scheffe’s) tests were applied to determine which groups differed from each other. Because the number of Caucasians, Aboriginals, and Others did not differ significantly between the High Intensity and Moderate Intensity groups, the two groups were combined for the purposes of these analyses.

There were no significant differences between Caucasians, Aboriginals, and Others on age, intelligence, antisocial personality disorder and general criminal risk score. However, there was a significant difference between Caucasians and Aboriginals on their SARA score.

Aboriginal offenders had significantly higher SARA scores than Caucasian offenders. There were no significant differences between the three groups on any of the pre-post treatment measures, except for the “sexual possessiveness” subscale of the Interpersonal Relationship Scale. Here, it was observed that offenders in the Other ethnicity group scored significantly higher than Caucasians on sexual possessiveness. It was also noted that there were some differences in the performance of ethnic groups on the results of the content quizzes administered at the end of each module (see below).

Parole Officer Feedback

The average number of months the parole officers supervised the offender was 10.03 ($SD = 6.55$). We attempted to ask parole officers about offenders’ attitudes and behaviour in the context of their current relationships; however, very few offenders (30%) were currently involved in a relationship. Of those who were, the majority of parole officers (71%) stated that they either had no concerns about the relationship or that the offender’s behaviour in the relationship had improved.

In addition to assessing offenders’ attitudes and behaviour in the context of relationships, we asked parole officers to judge offenders’ attitudes and behaviour post treatment on the items identified in the Goal Attainment Scale (GAS). Parole officers were asked to judge the GAS items on a 3-point scale (i.e., “does not appear to” exhibit the attitude or behaviour, “possibly/partially appears to,” “definitely appears to”). Many officers were unable to comment on changes since the commencement of treatment, since they were often unfamiliar with or could not recollect the offenders’ pre-treatment attitudes and behaviours. Table 2 illustrates the percentage of offenders post treatment who were rated in each of these categories on the various items. The parole officers’ judgments were primarily favourable, with the vast majority of parole officers observing at least some positive attitudes and behaviours on the GAS dimensions.

Parole officers reported that most offenders had not been charged or convicted of an offence since their participation in treatment (78%). Of the 22% who were charged or convicted since treatment, the type of charge/conviction varied, but none was spousal-related. Finally, parole officers were asked to make judgments about whether they felt the offender benefited from treatment. The majority of parole officers stated that they believed that the treatment was either somewhat or very effective (73%); 17% stated it was ineffective or somewhat ineffective and 9% were unable to say. Supplementary analyses were conducted to determine if any offender demographic or background variables were related to the parole officer perceptions of change and abusive behaviour. No

significant relationships were found, although the number of subjects for many of these analyses was low.

Table 2. Parole Officers' Judgments on Goal Attainment Scale Items – Post Treatment

| | Does not appear to | Possibly/partially appears to | Definitely appears to |
|--|---------------------------|--------------------------------------|------------------------------|
| Accept responsibility for his abusive/violent behaviour | 9.5% | 33.8% | 56.8% |
| Acknowledge his use of power and control tactics | 13.5% | 35.1% | 51.4% |
| Show empathy | 13.5% | 36.5% | 50.0% |
| Use appropriate problem solving skills | 13.5% | 41.9% | 44.6% |
| Use appropriate skills to manage and control his emotions | 12.2% | 37.8% | 50.0% |
| Use appropriate social skills | 10.8% | 36.5% | 52.7% |
| Recognize cognitive distortions | 12.5% | 50.0% | 37.8% |
| Minimize the consequences of his behaviour | 58.1% | 25.7% | 16.2% |
| Understand the dynamics surrounding his lifestyle | 9.5% | 23.0% | 67.6% |
| Understand his abusiveness pattern | 10.8% | 32.4% | 56.8% |

Note. n = 74.

Participant Feedback

At the end of every program, offenders completed a confidential feedback form on their opinion of the program. Offenders who completed the moderate program were asked: 1) Was the program interesting? (2) Did the program do a good job? (3) Would you be able to use the skills you have learned in the program in the community upon release? The results indicated that most offenders thought the program was very interesting (81%) and that the program did a very good job (84%). All (100%) of the moderate-intensity offenders who responded to the last question indicated that they would be able to apply the skills learned in the program in the community upon release.

Offenders who participated in the high-intensity program were asked to comment on the helpfulness of the program. Almost all participants (98%) stated that the program was helpful. Thus,

overall, participants in both the moderate- and high-intensity programs had positive comments about the specific program content. For more detailed information on the part of the programs that offenders found most and least helpful please refer to Kropp and Lee (2004).

Recidivism

Table 3 illustrates the recidivism rates for these three categories of offences for the treated and untreated groups in the moderate- and high-intensity programs.

Within the high-intensity group, there was a statistically significant difference between the treated and untreated offenders both for spousal violence ($p < .05$) and any violence ($p < .05$): 4% of the treated offenders committed a new spousal violence offence compared to 14% of the untreated offenders which is a 71 % decrease in spousal assault recidivism. Further, remaining with the high-intensity group, 11% of the treated and 26% of the untreated offenders committed a new violent offence, resulting in a reduction of violent recidivism of nearly 60%. For the moderate-intensity group, the same trends were observed but the differences were not statistically significant. However, when the moderate- and high-intensity groups were combined, once again there were statistically significant differences between the treated and untreated groups in the spousal violence and any violence ($X^2, p < .05$) categories; the results indicate a 69% reduction in spousal violence for the combined results of the two programs and a 47% reduction in any violence. The third category which included parole violations of any kind did not indicate a treatment effect.

Table 3. Recidivism Rates for Treated Versus Untreated Groups

| | Moderate Intensity | | High Intensity | | Combined Groups | |
|-------------------------|-----------------------------|-------------------------------|-----------------------------|-------------------------------|------------------------------|-------------------------------|
| | Treated (<i>n</i> = 76) | Untreated (<i>n</i> = 51) | Treated (<i>n</i> = 84) | Untreated (<i>n</i> = 35) | Treated (<i>n</i> = 160) | Untreated (<i>n</i> = 86) |
| Spousal Violence | 4% | 12% | 4%* | 14% | 4%* | 13% |
| Any Violence | 9% | 14% | 11%* | 26% | 10%* | 19% |
| Any Infraction | 24% | 26% | 24% | 37% | 24% | 30% |

Note. “Spousal Violence” is defined here as any actual, attempted or threatened violence towards a past or current intimate partner; “Any Violence” is violence (including threats) toward spouses and others; “Any Infraction” includes all criminal behaviour and conditional release violations not necessarily resulting in arrest.

* $p < .05$ (two-tailed); $N = 246$.

Table 4 expresses the recidivism data using odds ratios, which are measures of the relative likelihood of recidivism. For example, within the moderate-intensity group, untreated offenders were

3.25 times more likely than treated offenders to commit spousal violence, and 1.57 times more likely to commit any violence. Within the high-intensity group, untreated offenders were 4.5 times more likely than treated offenders to commit spousal violence, and 2.88 times more likely to commit any violence. Overall, when the moderate- and high-intensity groups are combined, untreated offenders were 3.76 times more likely than treated offenders to commit further spousal violence.

Table 4. Recidivism Odds Ratios for Treated Versus Untreated Groups

| | Moderate Intensity | High Intensity | Combined Groups |
|-------------------------|--------------------|--------------------|--------------------|
| Spousal Violence | 3.25 ($p = .09$) | 4.50 ($p = .03$) | 3.76 ($p = .01$) |
| Any Violence | 1.57 ($p = .43$) | 2.88 ($p = .04$) | 2.06 ($p = .06$) |
| Any Infraction | 1.10 ($p = .82$) | 1.89 ($p = .14$) | 1.39 ($p = .27$) |

Note. $N = 246$.

Table 5 presents the results of a binary logistic regression that tested how much the intensity level of treatment (moderate or high) or treatment completion in either program (treated versus untreated) contributed to the overall reductions in violent recidivism. The results suggest that family violence treatment, regardless of intensity level, had a significant effect upon reducing spousal violence ($OR = 3.83, p = .01$)⁸, in other words, this model confirmed the results in Table 4 indicating that offenders who complete either one of the programs are almost 4 times less likely to commit an act of spousal violence in the follow-up. It appears that treatment completion affected spousal violence recidivism in particular, but also generalized to reduce all violence ($OR = 2.18, p = .04$). The contribution of intensity level, however, was not significant. In fact, the intensity of treatment appears to be relatively unimportant. This could be interpreted as meaning treatment length and intensity do not necessarily matter. Alternatively, we know that the high-intensity group contained higher risk offenders. It could be, therefore, that the CSC has been successful at matching the risk and needs levels of offenders with the intensity of treatment.

⁸ The statistic used is the odds ratio (OR). The level of significance for the individual effects in this model is associated with their Wald statistics.

It is possible that, in the absence of this matching of risk and intensity level, the intensity level of treatment may have been important.

Table 5. Overall Effect (χ^2) and Relative Effects (*OR*) of Treatment Completion and Intensity on Recidivism

| | Overall Effect (χ^2) | Completion (<i>OR</i>) (Yes vs. No) | Intensity (<i>OR</i>) (Moderate vs. High) |
|-------------------------|-----------------------------|--|--|
| Spousal Violence | 6.74 ($p = .03$) | 3.82 ($p = .01$) | 1.10 ($p = .85$) |
| Any Violence | 4.93 ($p = .09$) | 2.18 ($p = .04$) | 1.58 ($p = .24$) |
| Any Infraction | 1.73 ($p = .42$) | 1.43 ($p = .24$) | 1.24 ($p = .47$) |

Note. Effects computed using logistic regression. $N = 246$.

Overall, the recidivism analyses suggest that the CSC high- and moderate-intensity programs are achieving the stated goal of reducing violence. The results are consistent with other sections of this report suggesting that treatment had a positive effect on the majority of the participants.

Factors Related to Success in Treatment

This section of the paper will look at which factors were linked to various criteria of successful outcome. We will report on factors that predict drop out, factors related to facilitators' final ratings of successful completion of the programs and factors associated with cessation of spousal abuse on release.

A correctional matrix and a series of t-tests examined the relationship of a number of factors the literature suggested would be related to drop out. We have a more extensive test battery for the participants of the high-intensity program so we were able to profile these offenders on more variables. Since the number of drop outs was low, however, we combined the moderate and high intensity program participants for most of the comparisons.

Correlational analyses as well as t-tests showed that ethnicity (Aboriginal or Other Ethnic), IQ, antisocial or borderline personality disorder diagnosis, spousal assault risk rating (SARA) and the extent of participants' substance abuse (either drugs or alcohol) were not related to drop out. The primary demographic difference between dropouts and completers was that

dropouts were significantly $t = 9.50, p < .05$ younger (dropouts: $M_{\text{age}} = 34.20$ vs. completers: $M_{\text{age}} = 37.05$). In addition, dropouts' overall criminal risk rating on the Statistical Information on Recidivism scale (Nuffield, 1982) was slightly, but significantly, higher, $t = -2.28, p < .05$.

Drop outs did not differ from completers on the pre-treatment self-report measures including measures of attitudes and jealousy. However, several dynamic (changeable) factors measured through facilitator ratings did distinguish the two groups. Some of the scales of the Family Violence Vignettes and most of the facilitator ratings on the pre-treatment Goal Attainment Scale (GAS) were significantly different for drop outs. In general, at the pre-treatment assessment, dropouts had been rated as treating their partners with less respect, being less responsive to treatment and having more negative attitudes and behaviours than those who completed treatment.

Since the intercorrelations on the items on the GAS were high and their relationship to drop out approximately the same, we selected one rating, the GAS item on motivation, for the regression equation. For the combined high and moderate intensity program participants, the variables of age and, risk and the motivation item on the GAS were entered into a logistic regression to assess the predictive accuracy of these variables in predicting who would drop out of the program. Since age did not add additional variance to a test model, the final model included only GAS motivation rating and overall risk. A test of the full model against a constant-only model was statistically significant [$\chi^2(2, 292) = 13.09, p < .001$], suggesting that the predictor variables (GAS motivation rating and overall risk), as a set, reliably distinguish between offenders who drop out and those who complete the program. The model classified 90% of the cases correctly. The variance in dropout accounted for was small with a Nagelkerke $\hat{R} = .09$, indicating that only 9% of the variability in dropout status was accounted for by the predictor variables. According to the Wald statistic, the SIR and GAS motivation items individually reliably predicted who would drop out (out $(-5.5, df 1, p < .05)$ and $5.03, df 1, p < .05$ respectively). These results indicate that the reasons for drop out are not well explained by the range of assessment measures and profiling information we used in the regression models (indeed, over 90% of the reasons for drop out were left unexplained by the current data). Nevertheless, we can, with the information we have, identify participants who are at higher *risk* for drop out and target them for more specialized additional intervention.

Information on additional variables assessing aspects of motivation was available for the high-intensity group. Analysis of this information found other variables associated with drop out: the interview-based rating by the facilitators that assesses motivation for treatment; the pre-treatment responsivity rating; and the total score on the Precontemplation stage on the URICA-DV (Levesque et al., 2000), a measure developed to assess the stage of change of participants with histories of domestic abuse. Risk rating as assessed on the SIR scale was not related to drop out for this high-intensity group. The final model that included the interview-based rating of motivation, the GAS motivation item and treatment responsivity rating was significant according to the Chi square statistic [$\chi^2(3, 292) = 8.88, p. <.05$], but the Wald statistics for all the individual items were not. The model classified 90.4% correctly but the model explained only 17% of the variance (Nagelkerke $R^2 = .17$). As was the case when we looked for the variables that predicted drop out in the combined moderate and high intensity programs, the variables correlated with drop out for the high intensity group participants could be used to identify offenders *at risk* for drop out, but most of the reasons why offenders drop out (83%) is not accounted for by the model.

The next treatment outcome variable we examined was overall progress on the goals of the program. This includes ratings of skill development, prosocial attitude change and development of coping strategies to deal with high-risk situations. An overall treatment score was computed based on the total score from all the post-treatment GAS items⁹.

We found that none of the demographic variables available on the total treatment population was related to outcome (age, ethnic group membership, overall criminal risk, spousal assault risk (SARA), and extent of substance abuse). Factors that were significantly correlated to outcome were pre-treatment ratings on the GAS and a number of the pre-treatment scores on the vignettes.

For the combined population of moderate- and high-intensity program participants, we placed the following factors in the regression model because they were most highly correlated with outcome: scores on the GAS pre-treatment items on acceptance of responsibility ($r = .49$), the initial motivation for treatment rating ($r = .42$), scores on the Abusive Relationships Inventory (Myths subscale), scores on the Family Violence Vignette on jealousy and scores on the Empathy Scale (Perspective Taking scale). This overall model explained 39% of the variance

($R^2_{adj} = .392$, $F_{4,218} = 36.12$, $p < 0.001$). The effect of each of the individual variables was also significant ($p < .001$). This confirms that offenders who start the program with more willingness to acknowledge the extent of their abusiveness, who are motivated for treatment, who have the ability to take the perspective of the victims, do not endorse abusive attitudes and have lower levels of jealousy are likely to continue to progress well in the program, completing the program with non abusive attitudes and a good level of skill development.

Additional variables were available for the high-intensity population and these were analyzed further. We found that Antisocial Personality Disorder (APD), IQ, the pre-treatment readiness-to-change score, pre-and -treatment interview ratings tapping offenders' motivation to change were related to outcome as were all of the ratings on the GAS. The variables were also inter-correlated. An initial model demonstrated that APD did not add additional variance. We chose three variables that were most highly correlated with outcome for the final regression model: treatment readiness, IQ and an interview question rating motivation at pre-treatment. Individually, each of the variables significantly contributed to treatment success, but together, they more accurately accounted for the final treatment outcome¹⁰. Essentially, the results indicate that brighter participants who were motivated for treatment learned more skills and had less abusive attitudes by the end of the program. The same factors that were unrelated to overall treatment progress for the combined group were also unrelated to overall treatment progress for the high-intensity group alone: overall risk for family violence (score on the SARA), ethnic group membership, extent of substance abuse, or age. In addition, Borderline Personality Disorder was not related to overall treatment progress.

The final outcome variable we examined was spousal violence on release. Once again, the base rates were low so we combined the results for both the moderate- and high-intensity program participants for a correlational analysis. Criminal risk rating (SIR scale score for non Aboriginals and general criminal risk rating for Aboriginal offenders), spousal violence risk rating (SARA), age, extent of substance abuse, ethnic group membership and all the scores on the pre-treatment battery were not significantly related to spousal violence on release. The only variables even weakly related to spousal violence on release were an interview based rating on the extent of abusive thinking patterns ($\rho = -.245$, $p < .08$) and Antisocial Personality

⁹ See Table 1 for a complete list of all the GAS items.

Disorder (APD) ($\rho = -.23$, $p < .05$). We were unable to run further analyses on these data because of the low N (i.e., the low number of cases of spousal violence who had been released at the time of the study).

¹⁰ This model contributed 35.7% of the variance ($R^2_{\text{adj}} = .357$, $p < .001$). The overall model was significant ($F_{4,82} = 14.57$, $p < 0.001$). All of the variables were individually significant ($p < .01$).

DISCUSSION

Overall, several lines of evidence provide encouraging support for the positive impact of both the high- and moderate-intensity programs. Intermediate results suggest that offenders are showing significant improvement on almost all the measures tapping factors related to intimate partner violence: reductions in attitudes in support of spousal abuse, improvement in skills related to conflict resolution and communication,, reductions in jealousy, and a better understanding of the factors related to their offending pattern and the development of plans to address these factors. Content quizzes indicated that most offenders understood and could apply the major concepts of the program. Parole officers' observations of offenders who had completed treatment provide support for the program, and the offenders themselves overwhelmingly stated that they found the program useful.

Preliminary analyses of recidivism demonstrated that treatment participants in the two programs were 69% less likely than untreated offenders to be involved in an incident related to spousal abuse in a 6-month follow-up on release from prison. Completion of the program also significantly reduced rates of general violence on release.

The program does not, however, appear to have an impact on general- criminality and suspensions related to technical or non technical violations. Profiles of most federal offenders who have incidents related to spousal abuse in their offence histories indicate that these offenders have histories of general criminality as well. Previous research on the typology of spousal abusers within the Canadian federal system has shown that only 13.5% can be classified as "spousal only/non-pathological" (Wexler, 2000). The program content focuses specifically on skills slinked to intimate relationships that would also apply to resolving or avoiding general interpersonal conflicts. There is no emphasis however, on avoidance of non-violent criminality. The results of this evaluation point to a need for additional intervention for spousal violence perpetrators with histories of non-violent criminality.

It is notable that some of the measures (i.e., Family Violence Vignettes, Treatment Readiness Scale, Goal Attainment Scale) that are *not* self-report in nature are most valuable for anticipating which offenders might drop out of treatment. Rather than being self-report, these measures are coded or rated by someone who knows the offender. A simple item in which facilitators rate the motivation of offenders for treatment and another item rating the offender on motivation to make changes related to his abusive behaviour were significantly related to drop out and to facilitators' final evaluations of progress. Previous research has demonstrated that staff working with offenders can make estimates of

the motivation of offenders to address their criminality that are related to recidivism (Stewart and Millson, 1995). Training on how to complete these scales is simple, and the scales themselves are completed in a couple of minutes if the rater is familiar with the offender and they are as strongly related to drop out as the URICA-DV, a longer measure that is completed by perpetrators and is more time consuming to train staff to score and interpret.

The current study did not find that any of these motivation ratings were significantly related to spousal violence on release. At this point in the study, the only variables related to spousal violence were antisocial personality and a general rating by the facilitators of the extent of the offenders' thinking and beliefs in support of violence against women. We expect that as the base rates increase with longer follow-up periods, we will be able to get a more thorough picture of the factors contributing specifically to intimate partner violence on release.

An evaluation design that includes drop outs as part of the comparison group is generally considered to be "naïve to the features of drop out." We do recognise the limitations of our decision to include drop-outs in our comparison group. Nevertheless, within CSC with its extensive menu of correctional programs and the legal mandate to address offenders' criminal behaviour through program participation, it is very difficult to identify a non-treatment comparison group. The majority of offenders in the comparison group did not complete the program for reasons that are not theoretically linked to outcome, however; that is, 40% were identified as needing the program, but for administrative reasons (e.g., the program was not offered prior to their parole dates, they were transferred to another institution, they did not speak the language of instruction was a conflict with work, etc.) were not able to start the program. Among the actual drop outs, 30% left the program for similar administrative reasons. Further research should apply methodology such as an instrumental variable analysis or a propensity scoring to adjust for any possible differences between the treatment and comparison groups that could not have been attributed to program treatment effect. Statistical methods like these have been used to improve quasi experimental evaluation design by controlling for potentially significant differences between the comparison and treatment groups (Jones, D'Agostino, Gondolf, & Heckert, 2001; Jones & Gondolf, 2002). In the future analyses, we will also be extending our follow-up period, although the 6-month period following release is generally acknowledged as the most critical time period for criminal recidivism. Likewise, offenders who

reoffend against their partners are most likely to do so within 6 months of program commencement (if they are receiving the program in the community; Gondolf, 2001a).

It is as illustrative to review what is not related to treatment outcome as what was related. Ethnicity was not related to any outcome variables including drop out, measures of recidivism or almost all the intermediate measures of treatment progress. This suggests that the program as it is designed and delivered is equally effective for different ethnic groups. There were, however, ethnic differences on the results of the quizzes. These results point to a need to review the quizzes to determine if they are appropriate for non Caucasians. We will be reviewing the language, literacy level, format and content of the questions in the quizzes.

In addition, program outcome was not related to Borderline Personality Disorder. Dutton's research had associated problematic attachment patterns related to Borderline disorders with spousal violence (Dutton, 1995) and had speculated that this group may have needed a differential treatment approach. The CSC programs examined in this study, however, appear to have benefited offenders identified with higher scores on Borderline symptomology as much as those with lower scores. Facilitators are trained in motivational techniques and to make reasonable adaptations for individual differences. They also provide individual sessions that supplement the group curriculum. These adaptations may allow the program to meet the responsivity principle by accommodating the cognitive and emotional styles of various program participants.

Previous studies had linked substance abuse to program attrition (Daly et al., 2001), a finding that was not confirmed in this research. It may be that when the program is delivered in a prison setting, drug and alcohol abuse do not seriously impinge on offenders' program participation to the extent that they might if the offenders were in the community on conditional release.

Future research plans will focus on identifying a less potentially biased comparison group and providing further differential analysis that will allow us to determine what components of the program have the most impact on reductions in recidivism. At this point, however, we are encouraged that, while we can improve on the design of the evaluation study, all of the variables associated with treatment impact currently point in a positive direction.

REFERENCES

- Andrews, D. & Bonta, J. (1998). *The Psychology of Criminal Conduct (2nd Edition)*. Cincinnati, Ohio: Anderson Publishing Co.
- Babcock, J. C., Green, C. E., & Robie, C. (2004). Does batterers' treatment work? A meta-analytic review of domestic violence treatment outcome research. *Clinical Psychology Review, 23*, 1023-1053.
- Bennett, L. & Williams, O. (last revised 2004). Controversies and Recent Studies of Batterer Intervention Program Effectiveness. Minnesota Center Against Violence and Abuse (MINCAVA). Applied Research Forum. National Electronic Network on Violence Against Women.
- Blackburn, R., & Fawcett, D. J. (1999a). The Antisocial Personality Questionnaire: An inventory for assessing personality deviation in offender populations. *European Journal of Psychological Assessment, 15*, 14-24.
- Blackburn, R. & Fawcett, D. J. (1999b). Manual for the Antisocial Personality Questionnaire (APQ). Unpublished Manuscript, University of Liverpool.
- Boer, D. P., Kroner, D. G., Wong, S. & Cadsky, O. (undated.) Assessing the dysfunctional beliefs of spouse assaulters: Development and validation of Abusive Relationships Inventory. (Unpublished manuscript).
- Correctional Service Canada (2001). Assessment manual for the moderate and high intensity family violence prevention programs. Ottawa, ON: Reintegration Programs Division. Revised 2006.
- Daly, J. E., Power, T. G., & Gondolf, E. W. (2001). Predictors of batterer program attendance. *Journal of Interpersonal Violence, 16 (10)*, 971- 991.
- Dobash, R. & Dobash, R. (2000). Evaluating criminal justice interventions for domestic violence. *Crime and Delinquency, 46(2)*, 252-270.
- Dowden, C. & Serin, R. (2001). Anger management programming for offenders: The impact of program performance measures. Research report R-106 2001. Ottawa, ON: Correctional Service of Canada.
- Dunford, F. (2000). The San Diego Navy Experiment: An assessment of interventions for men who assault their wives. *Journal of Consulting and Clinical Psychology, 68(3)*, 468-476.
- Dutton, D. G. & Hart, S. D. (1992). Risk markers for family violence in a federally incarcerated population. *International Journal of Law and Psychiatry, 15*, 101-112.

- Dutton, D. G. (1995). *The domestic assault of women*. Vancouver, BC: UBC Press.
- Dutton, D. G. & Haring, M. (1995), *Perpetrator Personality Effects on Post-Separation Victim Reactions in Abusive Relationships*, Vancouver, BC: UBC Press.
- Feder, L. & Forde, D. R. (2000). A test of the efficacy of court-mandated counseling for domestic violence offenders: The Broward experiment. (National Institute of Justice U.S. Department).
- Gendreau, P. (1996). The principles of effective intervention with offenders. In A. T. Harland (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply*, 117-130. California: Sage.
- Gendreau, P. & Andrews, D. A. (1996). *Correctional Program Assessment Inventory*. St. John, NB: University of New Brunswick.
- Gondolf, E. (2001a). Limitations of experimental evaluations of batterer programs. *Trauma, Violence, and Abuse*, 2, 79-88.
- Gondolf, E. (2001b). An extended follow-up of batterers and their partners. Final report. (1997-2001). Centers for Disease Control and Injury Prevention Research for Violence Against Women.
- Griffin, D. W. & Bartholomew, K. (1994). The metaphysics of measurement: The case of adult attachment. In K. Bartholomew & D. Perlman (Eds.), *Advances in personal relationships Vol. 5*: 17-52. London: Jessica Kingsley.
- Hanson, K. & Wallace-Capretta, S. (2000). Predicting Recidivism Among Male Batterers. Research Report 2000 – 06. Ottawa, ON: Department of the Solicitor General Canada.
- Hart, S. D., Kropp, P. R., Roesch, R., Ogloff, J. R. P., & Whitemore, K. (1994). Wife assault in community-resident offenders. *Canadian Journal of Criminology*, October, pp. 4.
- Hogue, T. E. (March, 1993). Goal Attainment Scaling: A measure of clinical impact and risk assessment. A Paper presented at the 3rd Annual Conference of the British Psychological Society. Division of Criminal and Legal Psychology, Harrogate, North York, England.
- Hupka, R. B. & Rusch, P. A. (2001). Interpersonal relationship scale. In B. F. Perlmutter, J. Touliatos. & G. W. Holden (Eds.), *Handbook of Family Measurement Techniques: Instruments & Index Vol. 3*). Thousand Oaks, CA: Sage Publications, pp. 148-150.

- Jones, A., D'Agostino, R., Gondolf, E., & Heckert, A. (2001). Assessing the effect of batterer program completion on reassault using propensity scores. Poster presented at the New Hampshire Conference.
- Jones, A. & Gondolf, E. (2002). Assessment of the effect of batterer program completion on reassault: an instrumental variables analysis. *Journal of Quantitative Criminology*, 18, 71-98.
- Kropp, R. P. & Lee, Z. (2004). Evaluation of the Correctional Service of Canada Family Violence Prevention Programs: Final report. B.C. Institute Against Family Violence.
- Levesque, D., Gelles, R. & Velicer, W. (2000). Development and validation of a stages of change measure for men in batterer treatment. *Cognitive Therapy and Research*, 24, 175-199.
- Nuffield, J. (1982). Parole decision making in Canada: Research towards decisions guidelines. Ottawa, Canada: Department of the Solicitor General Canada.
- Oldham, J., Clarkin, J., Applebaum, A., Carr, A., Kernberg, P., Lottermen, A. & Haas, G. (1985). A self-report instrument for Borderline Personality Organization. In T. H. McGlashan, ed., *The Borderline: Current Empirical Research. The Progress in Psychiatry Series*. Washington, D.C.: American Psychiatric Press, Inc., pp. 3-18.
- Paulhus, D. L. (1990). Measurement and control response bias. In J. P. Robinson, P. R. Shaver, and L. S. Wrightsman, eds., *Measurement of Personality and Social Psychological Attitudes*. San Diego: Academic Press. pp. 37-41.
- Serin, R., & Kennedy, S. (1997). Assessment Protocol for Treatment Readiness, Responsivity, and Gain: Guidelines for Clinician Ratings. Unpublished manuscript. Correctional Service of Canada, Research Branch.
- Shipley, W. C. (1967). *The Shipley Institute of Living Skills*. Western Psychological Services. Revised by Zachary, R. A. (1991).
- Stewart, L. & Millson, W. (1995). Offender motivation for treatment as a responsivity factor. *Forum on Corrections Research*, 7, 5-7.
- Taylor, B. G., Davis, R. C. & Maxwell, C. D. (2001). The effects of a group batterer treatment program in Brooklyn. *Justice Quarterly*, 18(1), 170-201.
- Wexler, A. (2000). Can typologies of male batterers be generalised to populations of federal inmates? Unpublished Masters Thesis. Carleton University.