



Men's Engagement in Gender-Based Violence Prevention

A Critical Review of Evaluation Approaches

PRINCIPAL INVESTIGATORS

Samantha Blostein, Eco-Ethonomics Inc.
Areetta Bridgemohan, Eco-Ethonomics Inc.
Ryan Turnbull, Eco-Ethonomics Inc.



CO-INVESTIGATORS

Humberto Carolo, Director of Programs, White Ribbon
Kate Bojin, Project Manager, White Ribbon

SUGGESTED CITATION:

White Ribbon Campaign. (2014). Men's Engagement in Gender-Based Violence Prevention: A Critical Review of Evaluation Approaches. White Ribbon Campaign, Toronto, Ontario. March 2014.

CONTACT

Humberto Carolo, Director of Programs, White Ribbon
365 Bloor East, Suite 203, Toronto, ON M4W 3L4
hcarolo@whiteribbon.ca | www.whiteribbon.ca
© White Ribbon Campaign, 2014

With funding support from Status of Women Canada



Status of Women
Canada

Condition féminine
Canada

TABLE OF CONTENTS

EXECUTIVE SUMMARY.....	4
1.0 CONTEXT FOR GENDER-BASED VIOLENCE PREVENTION PROGRAMS FOR MEN AND BOYS.....	5
2.0 BACKGROUND FOR GENDER-BASED VIOLENCE PREVENTION PROGRAM EVALUATIONS.....	6
2.1 Types of Entry Points.....	6
2.2 Types of Program Interventions.....	6
3.0 WHAT WE FOUND.....	10
3.1 Approaches	10
3.2 Data collection points	11
3.3 Evaluation design	11
3.4 Methods	11
3.5 Data analysis	12
4.0 WHAT CAN BE MEASURED?.....	14
4.1 Outcomes.....	14
4.2 Indicators.....	14
4.3 Capacity building for evaluation.....	14
5.0 CHALLENGES TO EVALUATION.....	16
5.1 Gaps in evaluations of gender-based violence prevention programming for boys and men.....	16
6.0 KEY LEARNINGS.....	20
7.0 CONCLUSION	21
REFERENCES.....	22
APPENDIX 1: SAMPLE OUTCOMES.....	25
APPENDIX 2: SAMPLE OUTCOME FRAMEWORK.....	27
APPENDIX 3: SAMPLE INDICATORS.....	28

EXECUTIVE SUMMARY

This literature review highlights key conceptual and evaluation approaches to engaging boys and men in gender-based violence prevention.

While limited research has been conducted regarding promising evaluative approaches, there are numerous promising research instruments being utilized, such as the Gender Equitable Men's Scale and "Violence Against Women and Girls: A Compendium of Monitoring and Evaluation Indicators" (USAID, 2008).

Complicating the issue of evaluation, there are diverse entry points and contexts in which boys and men are unpacking masculinity and gender-based violence; from workplace awareness initiatives to community-based workshops. Rigorous and long-term evaluation is essential to ensure that male engagement programming is impactful at multiple levels (individual, community and broader systems). Some of the challenges in the evaluation of male engagement programming are: participant self-selection, one off programming and lack of evaluation data across the long-term. The literature review also critically analyzes the term "engagement."

We suggest that engagement encompass a wide range of processes and activities to also include participation in social media forums (as oppose to solely physical presence in a workshop or campaign/initiative). While there has been progress in assessing attitudinal change however there is a lack of research which documents concrete behavioural change over the long-term. This can be due to lack of funding or capacity of projects to carry out effective evaluation of programs. It's clear that rigorous evaluations are needed to ensure the engagement of men and boys in primary prevention initiatives leads to a significant advance in the reduction of violence against women and girls. Unpacking masculinity from a diversity and critical approach is also important, as boys and men do not experience power in the same ways. Race, class, sexuality, ability, cultural and religious affiliation, age are important factors to consider within evaluation methodology and outcomes.

1.0 CONTEXT FOR GENDER-BASED VIOLENCE PREVENTION PROGRAMS FOR MEN AND BOYS

There is a global movement to engage men in initiatives to reduce gender-based violence (GBV) and increase awareness of the negative societal impacts of violence against women and girls. Violence against girls and women is a complex global challenge. To adequately address and prevent gender-based violence, this requires a critical analysis of its root causes and unpacking dominant notions of masculinity. This literature review will explore multiple types of interventions, evaluation approaches and outcomes which engage boys and men in violence prevention.

Violence against women, predominantly perpetrated by intimate male partners, is a widespread national concern. In Canada, half of all women have experienced at least one incident of physical or sexual violence since the age of 16 (Canadian Women's Foundation, 2011). Violence against women has many forms and includes: inter-personal or domestic relationships; sexual violence (rape and sexual assault); institutional, cultural and systemic forms of violence; and emerging forms of stalking and harassment using modern technology (Minerson et al., 2011). The wide-ranging forms of violence against women also pose a challenge in its monitoring and evaluation. There are groups of women that are at particularly high risk of gender-based violence including newcomer and Aboriginal women. Aboriginal women are eight times more likely to be killed by their intimate partner than non-Aboriginal women (Canadian Women's Foundation, 2013).

Meaningful, long-term evaluation data on Canadian interventions and programmes working with men and boys is severely lacking (Minerson et al., 2011; Wells et al., 2013). A detailed review of the literature found few formal evaluations for primary prevention strategies designed for initiatives aimed at working with boys and men before violence occurs. Due to its relatively new approach, most work to engage boys and men in ending violence is largely small scale and under-funded. To date, most evaluations have focused on secondary or tertiary methods, measuring the impact of programs with men after the use of violence. There is also a lack of legislative approaches or evidence-based policies that encourage the involvement of men in primary prevention of violence against women and girls (Wells et al., 2013).

In light of these issues, this literature review was undertaken to identify the type of evaluation work conducted on gender-based violence prevention programs, and common challenges experienced by these evaluations. White Ribbon staff recommended the eleven sources consulted for this review. Selected readings range from global, national, provincial and cross-cultural case studies, surveys, political and historical analyses, campaigns and evaluations. Key information was pulled from each source about interventions, indicators, outcomes, methodology, challenges and key learning for evaluating programming that engages men and boys in gender-based violence prevention.

2.0 BACKGROUND FOR GENDER-BASED VIOLENCE PREVENTION PROGRAM EVALUATIONS

2.1 TYPES OF ENTRY POINTS

From a review of the literature there are several types of programming or avenues to engage men and boys in the prevention of gender-based violence. All have the goal of changing social norms regarding men's behaviours and attitudes towards masculinity and violence against women. The "2012 Shift: The Project to End Domestic Violence" report published by the Calgary Area United Way, identified seven areas of promise or entry points for engaging Canadian boys and men in violence prevention (Wells et al., 2013). The report discussed these entry points specifically in relation to domestic violence but many programs have used similar entry points to engage men to address GBV generally and not exclusively between intimate partners. The entry points are:

- 1 **Engaging fathers**
- 2 **Men's health**
- 3 **The role of sports and recreation**
- 4 **The role of the workplace**
- 5 **The role of peer relationships**
- 6 **Men as allies**
- 7 **Aboriginal healing**

The authors of the Shift report define entry points as, "areas of opportunity or engagement where men and boys can become involved in learning, reflection and action that will increase their potential to be involved in healthy relationships and decrease the likelihood they will be abusive (Wells et al., 2013: 5)." Though there are countless programs to address violence, there are few that focus directly on engaging boys and men to stop first-time perpetration of violence (Wells et al., 2013). The areas listed above provide promising primary violence prevention entry points that can be leveraged and enhanced through effective program evaluation.

2.2 TYPES OF PROGRAM INTERVENTIONS

There were a range of programs whose evaluations were consulted for this review, these included: community-based, sports-based, health-based and workplace interventions.

Community-based interventions involve working with community members and leaders in awareness-raising educational workshops, community-wide campaigns and community-based programs aimed at gender equality. Sports-based interventions use recreational activities, games and sports tournaments as gathering points and spaces in which the construction of masculine norms can be assessed and analyzed. Coaches, athletes

and educators facilitate educational workshops and community campaigns focusing on engaging young and adult men to prevent violence against women and girls. Health-based interventions take place in health clinics, community centres and schools and address topics related to sexual and reproductive health (contraception and HIV/STIs) and violence (at school, against women, etc.). Health-based interventions involve capacity training for health professionals to facilitate educational workshops to engage men and boys to address health and gender-based violence. Workplace interventions such as staff leadership training address harassment in the workplace and gender equity at home and at work.

Narrow reach of interventions

One of the key recommendations from a systematic review of engaging boys and young men in the prevention of sexual violence was that the cultural intervention reach¹ of these interventions is relatively narrow (Ricardo et al., 2011). There is a need for interventions to go beyond “white males not at high risk of perpetration in US/Canada.” As noted by Lonsway (2009), “One of the most pressing needs in the field is thus to expand our efforts beyond schools and campuses into our wider communities and across age, gender, class, ability/disability, race/ethnicity, sexual orientation, etc.” (Ricardo et al., 2011).

Community-wide interventions

An example of a community-wide intervention took place in Uttar Pradesh, India. Panchayats, local leadership councils were targeted to engage men and women in the

rural villages through advocacy campaigns, youth groups, community outreach and 23 educational workshops. Though this effort 1,500 youth and men were taught about the negative impacts of gender-based violence and violence prevention tools. To gauge program impact over time, pre-, mid- and post-intervention surveys were conducted with group education participants. In-depth interviews with participants and facilitators were also conducted. To measure change in gender equitable attitudes, the Gender Equitable Men (GEM) Scale was completed by participants pre- and post- intervention. Interviews were conducted with female partners of participants to evaluate attitudinal changes as a result of the intervention. There were notable increases in self-reported gender equitable attitudes after the workshops, specifically in positive attitude changes regarding the division of childcare responsibilities and household relationships. There were also positive attitudinal changes in the area of improved communication with partners on gender equality and sexual pleasure and increased knowledge of laws against GBV and women’s property and abortion rights (Instituto Promundo, 2012).

Very few of the studies reviewed were **system-wide**². Rather, most were focused on specific, limited target populations. In a systematic review, four studies reviewed were characterized as system-wide. Three of these (CEDPA, 2011; Solórzano et al., 2008; Verma et al., 2008) showed significant changes in the outcomes of interest, including attitudes toward violence, attitudes toward gender and relationships with women, and use of violence against women (Ricardo et al., 2011). Another report recommended that program areas should also include: policy approaches and cross-sectoral approaches

¹ Cultural intervention reach refers to interventions engaging members from broad population groups, “especially those who are at higher risk of perpetration, and among target populations outside of the Global North” (Ricardo et al., 2011: 7).

² A ‘system-wide’ approach refers to interventions focused on broad-based populations vs. limited interventions targeted at a specific population (Ricardo et al., 2011).

(IPPF, 2010). These observations were underscored in one of the studies reviewed on fatherhood, whereby fathers suggested that they were connected to the cities in which they had chosen to have their family. Fathers were not separate from their communities and, in many ways, relied upon their communities for many aspects of the parenting process (White Ribbon Campaign, 2014a). This finding demonstrates the importance of the role

of the broader environment in shaping the behaviour of individuals and groups, and as such, program interventions should probably also orient themselves toward larger structural change. The implication for evaluations is that as there were fewer programs targeting broader change. There is also less documentation on effective evaluation practices of this larger change.

Defining and Deconstructing ‘Engagement’

Engagement in programming

Various reports stressed the importance of understanding the multi-faceted nature of male roles in gender relations, and the implications in developing effective engagement strategies (IPPF, 2010). It was also observed that it is necessary to articulate the benefits of engaging in this work more clearly to men and boys, and to develop a better understanding of these benefits among men and boys (IPPF, 2010). In another study, participants often reported ‘recruiting’ other fathers to attend the father-specific programming in order to extend the benefits to others (White Ribbon Campaign, 2014a).

In the literature reviewed, a recurring theme was the tension between mixed-gender vs. single-sex programming. It was noted in the systematic review, that although effectiveness results are varied and there is no clear answer on this question, one of the studies found that participants were more satisfied in mixed-gender groups, perhaps related to age (and cultural context) (Ricardo et al., 2011). The review concluded that participants may be more drawn to mixed-sex interventions.

An example of digital engagement

In Canada, the White Ribbon Campaign’s “It Starts With You, It Stays With Him” targets fathers through social marketing to motivate them to teach the boys in their lives about the importance of equal and healthy relationships with girls and women (Minerson et al., 2011; Wells et al., 2013). Interactive engagement tools such as a website, YouTube Channel, collection of nine digital stories and Facebook page are complemented with a discussion guide for use in the community and in the classroom and e-learning modules for community workers and primary and secondary school educators. An evaluation in 2009 indicated that after visiting the website 82% of men felt more prepared to positively influence the boys in their lives about healthy relationships and gender equality. It Starts With You has reached an estimated 1,501,371 people in Ontario including over 82,000 children and youth (White Ribbon Campaign, 2014b).

Engagement in evaluation

In various studies, engagement took the form of consultation of both men and boys, for whom the programs reviewed were mostly designed to involve. Others focused on the consultation of women, for whom those programs were mostly designed to benefit. The engagement took place in program development as well as the evaluation process (IPPF, 2010).

One of the theoretical frameworks described in a study of fatherhood, addressed the concept of 'involved fatherhood.' One commonly used theoretical framework in the literature identified three main components of father involvement: accessibility, engagement and responsibility (Lamb, 1981 cited in White Ribbon Campaign, 2014a). Engagement was described as the 'actual time spent with children engaged in activities that promote healthy child development' (Lamb, 2004 cited in White Ribbon Campaign, 2014a). It was noted that evidence of nurturing fathering practices is the most consistent predictor of gender equality (Coltrane, 1996 cited in White Ribbon Campaign, 2014a).

Challenges

As listed in The International Planned Parenthood Federation (IPPF) 2010 "Men Are Changing: Case study evidence on work with men and boys to promote gender equality and positive masculinities", common challenges faced by programs across the globe are:

1. Recruiting and retaining participants
2. Actively engaging the surrounding communities
3. Implementing efficient and accurate monitoring and evaluation systems
4. Ensuring that long lasting, sustainable programs are instituted, i.e. that effective GBV prevention with adult and young men have been taken up in large-scale institutions

In terms of barriers to individual men's engagement in Canada, the White Ribbon Campaign 2011 Issue Brief outlines:

1. Accountability Barriers - lack of men taking ownership to address violence against women
2. Awareness Barriers - men do not understand severity of issue in Canada
3. Privilege Barriers - blaming other men
4. Men's silence - about violence against women and girls

Broadening the definition of engagement

In looking through the literature for this review, no source had defined what is meant by engagement in relation to measuring men's involvement in violence prevention interventions. Establishing a working definition to evaluate men's engagement in programming is at the heart of what many groups are attempting to do. Without a working definition that includes important elements such as male responsibility and commitment, it is difficult to measure program effectiveness. With the rise and interest of social media, including Facebook, Twitter etc., this points to the need to expand our traditional notion of engagement to not only participation in physical spaces, but also virtual communities and networks.

3.0 WHAT WE FOUND

3.1 APPROACHES

Theoretical frameworks

In addition to the theoretical framework cited above to understand ‘father involvement,’ other commonly used theories underlying interventions aimed at primary prevention were: social learning theory³, social norms theory⁴, belief system theory⁵ and theories around bystanders⁶ (Ricardo et al., 2011). In another report, interventions were selected based on criteria from a World Health Organization review of 58 interventions engaging men and boys. Another framework that was cited, in order to understand the type of change being sought by interventions was the Gupta framework (IPPF, 2010). The Gupta framework consists of the following scale: gender-neutral, gender-sensitive, gender-transformative (IPPF, 2010). Gender neutral is defined as: no distinction between the needs of men and women, neither reinforcing nor questioning gender roles. Gender sensitive is defined as: recognizing gender norms but little attempt to transform them. Gender transformative is defined as: seeking to promote equitable relationships, change gender norms and social expectations (Minerson et al., 2011).

³ Social learning theory highlights principles of learning via modeling and stresses the point that perpetrators learn abusive behavior in their families of origin (Ricardo et al, 2011).

⁴ “According to social norms theory, people are often negatively influenced by inaccurate perceptions of how other members of their social group act or think” (Ricardo et al, 2011: 17). This means that a bystander’s willingness to intervene (in witnessing a violent act) can be increased by knowledge about the other person’s attitudes (Ricardo et al, 2011: 16).

⁵ Belief system theory asserts that intervention design must preserve people’s existing self-conceptions. Thus, rape prevention interventions that assume male participants are potential rapists are unlikely to achieve desired outcomes because men do not perceive themselves to be rapist (whether or not they have committed sexual assault) (Ricardo et al: 16).

⁶ The bystander approach, “attempts to influence men by appealing to beliefs they are shown to have about being potential helpers” (Ricardo et al, 2011: 16). This approach gives men a role in preventing violence including interrupting situations before or during an incident and speaking out against norms that lead to violence (Ricardo et al, 2011).

The gender transformative programs were found to be the most effective programs (IPPF, 2010). In one of the reports reviewed, it was recommended that programs develop clear conceptual frameworks, with a particular focus on clearly defining what is meant by gender equality – for both men and women (IPPF, 2010).

Participatory approaches

In some of the studies, it was highlighted that greater engagement of program participants and community members, in the program development and the evaluation led to greater success (IPPF, 2010). For example, Program H (Brazil), a promising initiative, focuses on the health of boys and young men working with “males aged 15 to 24 to address and redefine individual behaviors associated with harmful masculinity and social norms” (Wells et al., 2013: 16). Program H tried to “first and foremost to tap into the ‘alternative’ voices in low-income communities, that is, young and adult men who have been questioning traditional views of what it means to be a man. These voices of resistance to the dominant versions of masculinity helped us to develop a set of objectives (what we expect or hope from young men after their participation in the initiative) and to develop an evaluation methodology” (Barker et al, 2007: 1).

To serve as a benchmark for the Program H evaluation, those with ‘alternative voices’ were consulted in focus groups and pre-testing to develop indicators based on a deep understanding of locally held beliefs and values. Desired outcomes were identified through an ongoing

interaction and discussion with a group of young men who acted as advisers and peer promoters to the researchers. Adult and young women in the communities were consulted and affirmed that these outcomes were the attitudes they wanted from men (Barker et al, 2007; Instituto Promundo, 2008, 2012).

A number of studies underscored the importance of including women in the evaluation of the effectiveness of the intervention, although women were not considered the primary focus of the intervention. For example, in a case study described in an IPPF (2010) report, women involved in program design and in the evaluation process were asked to comment on changes they observed in program participants.

3.2 DATA COLLECTION POINTS

The importance of baseline data

A number of studies highlighted the importance of collecting comprehensive baseline data (IPPF, 2010; Harris-Decima, 2010; White Ribbon, 2012). The collection of this baseline data was important for program design purposes as well as measuring changes in the population being targeted. In one case study, extensive baseline data allowed program designers to establish priorities for target groups, which in turn relates to the differential analysis point, described in more detail later in this review. Another interesting example was the use of a street survey to collect baseline data for the program (IPPF, 2010).

3.3 EVALUATION DESIGN

Experimental, quasi-experimental and descriptive methods

There were a number of different designs described in the reports reviewed. There were quite a number that had no control conditions, others that had alternative interventions (i.e. different treatment conditions), others that used delayed interventions (for e.g. one intervention group would start and the next would start after a delay or the intervention took place in phases), others compared the treatment group with an unrelated intervention. In a systematic review focusing on rape prevention interventions, there were very few randomised controlled trials; more cluster randomized studies and the majority were quasi-experimental studies with comparison groups (Ricardo, 2011). There was an agreement in the literature reviewed, that there is a need for more rigorous evaluation designs, understood as experimental designs, to strengthen the body of evidence available on gender-based violence prevention efforts. It was recommended that spaces should be built into the program for reflection, particularly when dealing with programs that are being expanded, to generate insight about programme implementation (IPPF, 2010).

3.4 METHODS

Qualitative and quantitative

The majority of the evaluations drew upon mixed methods, that is, both qualitative and quantitative methods. This enhanced the ability of program evaluations to triangulate data (particularly with regard to self-reporting), to investigate why certain outcomes were observed, as well as to understand the impacts of process measures on outcomes. For example, in an evaluation

of an intervention aiming to decrease violence against women through changing attitudes, it was found that there was a statistically significant self-reported decrease in the use of violence against female partners. Qualitative results further affirmed that the group education and campaign activities led to increased discussion amongst men about gender equality and decreased support for attitudes encouraging men's use of intimate partner violence (Instituto Promundo, 2012).

Instruments

A number of scales were commonly cited in the literature. These included: Illinois Rape Myth Acceptance Scale, Burt Rape Myth Acceptance Scale, Attitudes toward Women Scale, Conflict Tactics Scale, Gender Equitable Men Scale (GEM), Rape Empathy Scale and the Attraction to Aggression scale (to identify high-risk males) (Ricardo et al., 2011). These more frequently used evaluation tools may be most useful in terms of generalizability (Ricardo et al., 2011).

3.5 DATA ANALYSIS

In many evaluations, *differential analyses* by sub-group were conducted. This addresses larger equity concerns that relate to this work. Even when working with general populations, there is a need to increase the practice of analyzing effects on sub-groups with higher baseline risk of perpetrating sexual violence in terms of attitudes and other relevant indicators. Those identified as “high risk” may have experienced sexual violence or already perpetrated forms of sexual violence (Ricardo et al., 2011). Especially when baseline risk varies more, or when working with a more heterogeneous group in general, differential effectiveness analysis is critical to understanding which interventions are most effective

with which populations. The literature noted some key populations that need to be identified in program effectiveness analysis. These groups include: socio-demographic data (age, gender, income, educational level); rural vs. urban; French vs. English; Aboriginal vs. non-Aboriginal; experience of violence as a child and later perpetration of violence (due to their potentially high correlation); high-risk vs. low-risk men and women (Ricardo et al., 2011). Another report stated that for some of the evaluations reviewed, there was a lack of *disaggregation by process indicators*, meaning that data is not consistently separated by group (such as male sex workers and men who have sex with men). This data could provide important information to practitioners (for e.g. are groups facilitated by professionals or peers more effective) (IPPF, 2010).

Gender Equitable Men Scale (GEM Scale)

One of the scales worth highlighting due to its common usage is the Gender Equitable Men Scale (GEM Scale). It is a psychometric attitude scale, administered by survey to male respondents in which they self-rate their attitude to statements based on degree of agreement. The GEM Scale has been validated in more than 20 sites around the world. It is useful in pre- and post-tests to observe change in how men answer the questions before/after an intervention. Depending on intervention setting, there are often 20-30 attitudinal statements regarding gender roles pertaining to the following areas:

- In the home and child-rearing,
- Sexual and reproductive health, sexual relationships HIV/STI prevention, and
- Norms about partner violence, and homophobia (Instituto Promundo, 2012).

The GEM Scale that measures changes in men's attitudes toward gender norms, was tested, and validated in the Instituto Promundo Program H gender-based violence prevention program. The following is a list of indicators from that research (Pulerwitz and Barker for Instituto Promundo, 2008):

Factor 1: Inequitable Gender Norms

1. It is the man who decides what type of sex to have.
2. A woman's most important role is to take care of her home and cook for her family.
3. Men need sex more than women do.
4. You don't talk about sex, you just do it.
5. Women who carry condoms on them are "easy."
6. A man needs other women, even if things with his wife are fine.
7. There are times when a woman deserves to be beaten.
8. Changing diapers, giving the kids a bath, and feeding the kids are the mother's responsibility.
9. It is a woman's responsibility to avoid getting pregnant.
10. A man should have the final word about decisions in his home.
11. Men are always ready to have sex.

Factor 2: Equitable Gender Norms

1. A couple should decide together if they want to have children.
2. In my opinion, a woman can suggest using condoms just like a man can.
3. If a guy gets a woman pregnant, the child is the responsibility of both.
4. A man should know what his partner likes during sex.
5. It is important that a father is present in the lives of his children, even if he is no longer with the mother.
6. A man and a woman should decide together what type of contraceptive to use.
7. It is important to have a male friend that you can talk about your problems with.

When using the GEM Scale to assess men's attitudes it is important to include questions addressing variables that are related to gender norms such as socio-demographic status, current safer sex behaviour, and relationship history of physical violence (Barker and Pulerwitz, 2008).

4.0 WHAT CAN BE MEASURED?

4.1 OUTCOMES

See Appendix 1 for list of selected outcomes. Appendix 2 illustrates a sample outcome framework that may be useful for considering outcomes at various levels of intervention and impact.

Various reports highlighted the over-reliance on attitudinal changes in program evaluations (Ricardo et al., 2011). Changes in attitudes have been linked to improvements in non-violent behavior outcomes in the research literature. However, there is little evidence of the effectiveness of interventions to actually decrease boys' and young men's perpetration of violent behaviors in the long-term (Ricardo et al., 2011). Over-reliance on the use of attitude measures as proxies for behaviors is therefore problematic. Attitude may have a key role in promoting individual and social change more broadly. However attitude change alone is not a sufficient outcome for gender-based violence prevention efforts (Ricardo et al., 2011).

4.2 INDICATORS

There were many excellent sources for indicators in the literature reviewed. One such resource is the: "Violence Against Women and Girls: A Compendium of Monitoring and Evaluation Indicators" compiled by USAID East Africa in 2008. Please see Appendix 3 for a list of selected indicators for evaluating programming targeting various groups.

4.3 CAPACITY BUILDING FOR EVALUATION

There was limited in-depth discussion of capacity building for evaluation in the resources reviewed. There were recommendations in reports, given some instances of inconsistent quality of monitoring and evaluation data, for greater training of staff in monitoring and evaluation (ME) data collection methods as well as training in the importance of this data (IPPF, 2010). Some of the intervention evaluations reviewed did include training of staff in data collection and utilization of a common evaluation framework as well as training of management focused on ME. One such example was an intervention which included the following capacity building activities: follow-up training; helping project staff develop monitoring tools and discuss strategies for developing the project; a 2-day staff training for team building, identifying successes/challenges; increasing capacity for training and recommendations for program development (IPPF, 2010).

In a large study undertaken assessing attitudes towards gender violence, there was an extensive discussion of the training provided to interviewers, in order to minimize the bias in the data collection process. A detailed training session was undertaken with all interviewers selected, including the supervisor who was assigned to monitor the interviewing team. Some of the topics covered included:

- ▶ An explanation of the survey objectives;
- ▶ An explanation of the sensitivities associated with the study

- ▶ Specific techniques designed to encourage honest reporting, including techniques for developing rapport, instilling trust in the respondent, and assuring him or her of confidentiality;
- ▶ A review of the survey design, including termination points, skip patterns, and question wording;
- ▶ A question and answer session;
- ▶ Pre-test of the survey instrument; and
- ▶ A debriefing to discuss any problems and obtain feedback from the interviewers on the survey instrument (Harris-Decima, 2010).

5.0 CHALLENGES TO EVALUATION

5.1 GAPS IN EVALUATIONS OF GENDER-BASED VIOLENCE PREVENTION PROGRAMMING FOR BOYS AND MEN

One of the recurring gaps stated in the reports reviewed, was the over-reliance of evaluations on self-reported data. The exclusive reliance on self-reports makes the results susceptible to social desirability bias (i.e. the pressure for the respondent to provide certain answers). One of the studies, a telephone survey of attitudes, included an extensive reflection on the factors that affect respondent honesty as well as mitigating strategies. This study noted: “There are a number of factors that affect respondent honesty when self-reporting behaviour and attitudes, including:

- ▶ Mode of data collection;
- ▶ Characteristics of the interviewer;
- ▶ Interviewer/respondent interaction;
- ▶ Socio-demographic characteristics of the respondents;
- ▶ Presence of third parties; and
- ▶ Extent of the perceived threat from answering questions honestly with risk.”

The study took this information into account when designing their methodology and training of the data collectors. The report states that “more honest reporting” generally occurs when the respondents’ characteristics

such as age, race and gender, are similar to those of the interviewer (Harris-Decima, 2009: 10). The report also contends that honest reporting is fostered when the respondent has assurance of confidentiality and when interviewers maintain a neutral and unbiased tone regardless of the respondent’s responses (Harris-Decima, 2009).

The following is listed in the New Brunswick Attitudinal Survey on Violence against Women, and was implemented by the Harris/Decima research team to ensure more honest reporting in the data collection process (Harris-Decima, 2009):

- ▶ Stressing that participation in the study is voluntary;
- ▶ Providing persuasive assurances of confidentiality;
- ▶ Assigning male interviewers to survey male respondents, and female interviewers to survey female respondents, where possible;
- ▶ Encouraging respondents to voice their views throughout the course of the interview;
- ▶ Providing respondents with assurances that there are no right or wrong answers;
- ▶ Ensuring that interviewers convey to respondents that all answers are acceptable;

- ▶ Implementing practices to ensure that no third parties are present during the course of the interview, or are able to overhear the interview; and
- ▶ Providing respondents with the survey registration number and telephone number for contacting the Marketing Research and Intelligence Association (MRIA) to receive a confirmation about the nature of the study and its anonymity (Harris-Decima, 2009: 11).

Furthermore, the study investigators trained data collectors to adopt specific techniques to encourage honest reporting (including techniques for developing rapport, instilling trust in the respondent, and assuring the respondent of confidentiality) (Harris-Decima, 2009).

This research was a comprehensive study of attitudes towards violence against women in New Brunswick. However, although prevalence rates and other research related to behaviour were included at the beginning of the report, the results were reported without reference to actual behaviour. This is a critique made of evaluations in this area. In a systematic review, the authors conclude, as it relates to rape prevention efforts: “there is an over-reliance on the use of attitude measures as proxies for behaviors; While attitudes have an important role to play in promoting individual and broader social change, it is necessary for researchers and practitioners to move beyond the assumption that attitude change in and of itself is a sufficient outcome for rape prevention efforts. ***The fact that there are many interventions that have demonstrated impact on attitudes correlated to violence is a promising indication that programs are moving in a***

positive direction. However, it is not sufficient, and there is a need to measure behaviors and actual rates of sexual violence” (Ricardo et al., 2011).

Not only was it noted that there was an over-reliance on self-reporting in evaluations, but also that in many of the programs, ***participants were self-selected***. This is understandable in terms of outreach efforts, ethics and even in terms of the implications of voluntary participation in attitude and behaviour change. However, evaluation results for programs with self-selected participation, should be interpreted within the lens that those who self-select to participate may be those most motivated to change (Stephens & George, 2009; from systematic review). As a result, positive outcomes with general groups may overestimate prevention effectiveness (Ricardo et al., 2011). An interesting approach that was documented in one of the cases reviewed, was the interview of 11 participants who dropped out of the program (IPPF, 2010). This is significant because it sought feedback to understand why the participants left and how the program could have been better suited to meet their needs. Very few of the programs reviewed mentioned intentionally attempting to capture the perspectives of participants who dropped out of the program. This enables the program to access potentially valuable information about how to retain participants and prevent dropout.

Challenges in comparing program effectiveness

It was noted that there is a wide variety of outcome measures (and evaluation tools), making comparisons of program effectiveness difficult. For example, in a systematic review, it was noted that there were a wide range of evaluation tools – 96 in total in the studies reviewed (Ricardo et al., 2011). This same review

observes: “The benefits of identifying a valid and reliable set of outcomes that can be used to evaluate the effectiveness of a wide range of sexual assault prevention programs would be of enormous benefit to the field, where widely varying outcome measures makes comparisons between evaluated programs problematic” (Ricardo et al., 2011). The measures that have been utilized in multiple studies and have been adequately tested for reliability and validity may provide a good starting point. It would be helpful if researchers working in the field could come together to agree on a standard set of measures that could be utilized as often as possible and appropriate (Ricardo et al., 2011). It may also be helpful to encourage use of a narrow range of indicators that work well across cultural settings. Again, the measures that have been utilized in multiple studies and have been adequately tested for reliability and validity may provide a good starting point for identifying solid measures, encouraging their use, and developing other measures to fill gaps where solid measures have not yet been developed or adequately tested (Ricardo et al., 2011).

Limitations of one-off programming

Practitioners need to know how much of an intervention is necessary in order to achieve the desired outcomes, while not wasting resources by providing more services than are necessary. The findings from this review do not provide a definitive answer to this question, in part because most interventions were not tested at multiple dosages. One study (Banyard et al 2007) that was reviewed tested the effects on non-sexual violence of two different levels of an intervention – a one-session intervention and a three-session intervention. Participating in both the one-session and three-session intervention produced significant changes, however the *group that received the lengthier program (three*

sessions) showed a more significant increase in positive bystander attitudes and lower rape myth acceptance than participants in the one-session group. Clearly, decisions about the extent of how often boys and men are engaged in programming should not be made based on the results of this one study. However, these findings suggest that additional research regarding how often boys and men participate in GBV interventions may produce useful findings for the field (Ricardo et al., 2011).

Lack of longer-term follow-up

A number of reports reviewed cited the lack of longer periods of follow-up, and of analyses examining differential levels of sustained change across participants. In a systematic review it was noted that outcomes were measured at a wide range of time points, from immediately after the intervention to four years post-intervention. This provides some information on both short and long-term effects. Most outcomes were measured immediately or within a short time period. A great majority of the studies reviewed did not follow participants for more than a few weeks, with 35% (n=23) having no follow-up beyond an immediate post-test, and another 17% (n=11) following and testing participants only 1-3 weeks after intervention. While it is understandable that resources are limited, stifling the ability for long-term follow up, this has negative implications in assessing long-term outcomes and program effectiveness.

There needs to be a sufficient length of time between assessments for program participants to have had the opportunity to engage in the behaviors of interest, and for their beliefs, attitudes, and knowledge to have grown, shifted, and changed. As an example, Foshee et al. (2004) found some significant changes in behaviors

that were detected only at the four-year follow-up. The lack of longer-term follow-up is a serious limitation in the studies overall. It was found that there was a lack of reporting about outcomes for more than one year impact (Ricardo et al., 2011). Another example of the importance of timeline of impact is in length of time it took to achieve certain outcomes in a school-based education program. A change was noted in the report of number of incidents of physical violence and use of passive aggressive strategies - in year 2 for girls, but in year 3 for boys (IPPF, 2010).

Pre-test issues

Additionally, there are concerns related to the potential impact of pre-tests. For example, Foubert and Marriott (1997) note a concern that administering the Burt Rape Myth Acceptance Scale raises awareness among participants. Lonsway and Kothari (2000) also talk about issue of pre-tests, stating that: “this problem is exacerbated when a pre-test is used, because it trains participants in exactly how to provide the right answers. In fact, several studies have documented positive effects that are apparently due to pre-test assessment (i.e., sensitization effects), when scores of pre-tested participants are compared with those who were not exposed to a pre-test (for a review, see Breitenbecher, 2000). Therefore, it is best not to use a pre-test-post-test design with only a single group of participants. Without a control group of individuals who did not participate in the program, the findings from this type of research cannot be interpreted” (Cook & Campbell, 1979).

Need for more rigorous evaluation

A critical finding in the systematic review read for this report is that there is a *need for more rigorous evaluation designs*, particularly in terms of randomization. However, random assignment is frequently challenging in the

real world, for several reasons. First, programs tend to be administered using intact groups (for e.g. classes, sports teams, fraternities, sororities, dorm floors) rather than individuals who can be randomly assigned to one condition or the other. It is often difficult or impossible to generate these groups randomly because they have already been created or because of scheduling difficulties. Second, organizations such as schools and community centers are often reluctant to randomly select some participants to receive a potentially helpful intervention, while denying this opportunity to others (Jaycox et al 2006; Flay and Collins 2005).

One alternative to randomization at the individual level is the cluster-randomized trial, in which schools, organizations, or communities are matched and randomly assigned to a treatment or control in pairs. While this is a possible solution when individual level randomization is not possible, it is not ideal because it requires much greater sample sizes in order to achieve statistical power needed to detect significant effects, and because of potential bias introduced by having more similar participants within a cluster, such as a school or community (Flay and Collins 2005; Murray 1998). Use of the cluster-randomized design requires more advanced and rigorous statistical analyses, which require resources that are sometimes beyond the capacity of evaluation projects. Future studies may need to utilize individual level random sampling when possible, increase the number of clusters, and utilize more sophisticated statistical methods to account for intra-class correlation issues and other problems associated with a small number of clusters (Ricardo et al., 2011).

6.0 KEY LEARNINGS

Below, key recommendations for evaluations of gender-based violence prevention programming are listed. These stem from the observations of the challenges and gaps that exist within gender-based violence programming evaluation.

- 1 Balance attitudinal change with behavioural change measures.
- 2 Triangulate self-reported data with other measures.
- 3 Develop a standard set of measures.
- 4 Conduct further research on engagement.
- 5 Differential analysis of program effectiveness should be conducted as it relates to high-risk subgroups, as well as timeline of program impact.
- 6 Build in longer follow-up periods measuring program impact.
- 7 Expand the reach of interventions
- 8 Value of randomized design evaluations.
- 9 Include research on costs associated with programming. This will enable funders to make working with boys and men a priority knowing the costs associated (IPPF, 2010).
- 10 Evaluations should consider issues of scale-up when evaluating pilot programs.
- 11 Use culturally appropriate and gender sensitive methods for intervention and evaluation.
- 12 Further research on capacity building for evaluation in gender-based violence prevention programming.
- 13 Consider how often boys and men participate in the intervention when measuring effectiveness

7.0 CONCLUSION

The wide-ranging forms of violence that women and girls experience in Canada present a challenge to monitor and evaluate programs aimed at engaging boys and men. This review of the literature served to assess evaluations for programs and interventions, which engage boys and men in addressing gender-based violence. The literature review highlights gaps in defining the term engagement and a lack of evaluation approaches focused on primary prevention strategies. The limitations of self-reporting point to the need to address bias and gaging actual attitudes and behaviours, as oppose to participants feeling pressured to answer in socially appropriate ways (i.e. condoning rape and violence against women). In addition, the need for long-term, systematic evaluation is also critical to show how attitudinal change can impact behavioural changes throughout boys and men's lives at multiple levels of analysis (individual, community and systemic levels).

Based on the literature reviewed, there are a multitude of programs that are attempting to address this issue, namely through entry points in community, sports, and health-based interventions. Gender transformative programs which seek to promote equitable relationships and change gender norms have been found to be the most effective interventions. On the other hand, effectiveness results are varied between mixed-gender versus single-sex programming. This is an important area to consider in various levels of program design and evaluation. Importantly, many programs have proven to impact attitudes correlated to violence. This is a promising sign that interventions are making a positive difference. Specific research instruments and tools were explored, such as the GEM scale and numerous scales which assess rape myth acceptance, however a critical approach should be applied to assess boys and men's support for girls and women's rights to not only survive gender-based violence, but to thrive in both public and private spheres. Encouragingly male engagement work continues to be scaled up, however a critical analysis must be applied to ensure participation is meaningful and impactful, ultimately to promote the human rights of girls and women- in all their diversity.

REFERENCES

Banyard, V. L., Moynihan, M. M., Plante, E. G. (2007). Sexual violence prevention through bystander education: An experimental evaluation. *Journal of Community Psychology*, 35 (4), 463–481.

Barker, G., Ricardo, C., Nascimento, M & Pulerwitz, J. (2007). How do we know if men have changed? Engaging men and boys in changing gender-based inequity in health: Evidence from programme interventions. World Health Organization.

Bloom, S. (2008). Violence against women and girls: A compendium of monitoring and evaluation indicators. MEASURE Evaluation for USAID East Africa Regional Mission and the Interagency Gender Working Group.

Breitenbecher, K. H. (2000). Sexual assault on college campuses: Is an ounce of prevention enough? *Applied & Preventive Psychology*, 9, 23–52.

Canadian Women's Foundation. (2013). The Facts About Violence Against Women. Retrieved December 11, 2013 from <http://www.canadianwomen.org/facts-about-violence>

CEDPA (2011). UDAAN: Towards a Better Future Final Evaluation Report. Unpublished. CEDPA: Washington, DC.

Cook, T.D. and Campbell, D.T. (1979). *Quasi-Experimentation: Design and Analysis for Field Settings*. Rand McNally, Chicago, Illinois.

Flay B. R & Collins, L. M. (2005). Standards of evidence: criteria for efficacy, effectiveness and dissemination. *Prevention Science*, 6:151–175.

Foshee, V. A., Bauman, K. E., Ennett, S. T., Linder, G. F., Benefield, T. & Suchindran, C. (2004). Assessing the long-term effects of the Safe Dates program and a booster in preventing and reducing adolescent dating violence victimization and perpetration. *American Journal of Public Health*, 94, 619–624.

Foubert, J. D. & Marriott, K. A. (1997). Effects of a sexual assault peer education program on men's belief in rape myths. *Sex Roles*, 36, 257-266.

Harris/Decima. (2009). Attitudinal survey on violence against women. Detailed Report. Prepared for the Province of New Brunswick Executive Council Office Women's Issues Branch.

Instituto Promundo (2012). Engaging men to prevent gender-based violence: A multi-country intervention and impact evaluation study.

Jaycox L. H, McCaffrey D, Eiseman E, Aronoff J, Shelley GA, Collins R. L, Marshall G. N. (2006). Impact of a school based dating violence prevention program among Latino teens: a randomized controlled effectiveness trial. *J Adolesc Health*; 39(5): 694-704.

Lonsway, K. A., Kothari, C (2000). First Year Campus Acquaintance Rape Education: Evaluating the Impact of a Mandatory Intervention. *Psychology of Women Quarterly*, 24; 220-232. Cambridge University Press. Copyright 2000, Division 35, American Psychological Association.

Lonsway, K. (2009). Rape prevention and risk reduction: Review of the research for practitioners. Harrisburg, PA: VAWnet, a project of the National Resource Center on Domestic Violence/Pennsylvania Coalition Against Domestic Violence. Retrieved from <http://www.vawnet.org>

Minerson, T., Carolo, H, Dinner, T., & Jones, C. (2011). Issue Brief: Engaging men and boys to reduce and prevent gender-based violence. Status of Women Canada.

Murray, D. M. (1998). Design and analysis of group-randomized trials. Oxford, UK: Oxford University Press.

Pulerwitz, J. and Barker, G. (2008). Measuring attitudes toward gender norms among young men in Brazil development and psychometric - Evaluation of the GEM Scan. *Men and Masculinities*, (10): 3.

Ricardo, C., Eads M., & Barker, G. (2011). Engaging boys and young men in the prevention of sexual violence: A systematic and global review of evaluated interventions. Sexual Violence Research Initiative and Promundo.

Solorzano, I., Bank, A., Pena, R., Espinoza, H., Ellsberg, M., Pulerwitz, J. (2008). Catalyzing Personal and Social Change Around Gender, sexuality, and HIV: Impact Evaluation of Puntos de Encuentro's communication Strategy in Nicaragua. Puntos de Encuentro. CIDS/UNAN Leon, PATH, and Horizons Program.

The International Planned Parenthood Federation (IPPF). (2010). Men Are Changing: Case study evidence on work with men and boys to promote gender equality and positive masculinities. London.

Verma, R., J. Pulerwitz, V. S. Mahendra, S. Khandekar, A. K. Singh, S. S. Das, S. Mehra, A. Nura, and G. Barker. (2008). Promoting gender equity as a strategy to reduce HIV risk and gender-based violence among young men in India. Horizons Final Report. Washington, DC: Population Council.

Wells, L., Lorenzetti, L, Carolo, H, T. Dinner, T. Jones, C. Minerson, T& Esina, E. (2013). Engaging men and boys in domestic violence prevention: Opportunities and promising approaches. The University of Calgary, Shift: The Project to End Domestic Violence.

White Ribbon Campaign. (2012). Men's Attitudes and Behaviours Toward Violence Against Women Report. Ontario Men's Survey (unpublished).

White Ribbon Campaign. (2014a). Give Love, Get Love. The Involved Fatherhood and Gender Equity Project. White Ribbon Campaign

White Ribbon Campaign, (2014b). Ontario Women's Directorate Interim Report. White Ribbon Campaign (unpublished)

APPENDIX 1:

SAMPLE OUTCOMES	
Men's Awareness	<ul style="list-style-type: none"> ▶ Awareness of risk taking behaviours, safe behaviours (IPPF, 2010); ▶ Awareness of policy makers of concerns of MSM and inclusion of their issues in a national HIV/AIDS agenda (IPPF, 2010); ▶ Beliefs about accuracy of TV representation of violence in real life; recognizing stereotypes and hidden messages in magazine ads (IPPF, 2010).
Men's attitudes and efficacy	<ul style="list-style-type: none"> ▶ Men's perceptions and interest in women's health (IPPF, 2010); ▶ Attitudes towards dating violence and knowledge about forms of abuse; attitudes towards dating violence (IPPF, 2010); ▶ Double standard of tolerance for violence by girls than violence by boys; ▶ Development of positive norms (e.g. responsibility and caring for one's sexual partners) (IPPF, 2010); ▶ Support for inequitable gender norms and sexual harassment of girls and women (IPPF, 2010); ▶ Intention/likelihood to perpetrate rape or other forms of sexual violence (Ricardo et al., 2011); ▶ Intention/likelihood to perpetrate non-sexual forms of violence against an intimate partner (Ricardo et al., 2011); ▶ Rape-supportive attitudes, including rape myth acceptance (Ricardo et al., 2011); ▶ Attitudes towards gender-based violence (Ricardo et al., 2011); ▶ Attitudes towards intimate partner violence (Ricardo et al., 2011); ▶ Attitudes towards interpersonal violence (Ricardo et al., 2011); ▶ Empathy for rape or sexual assault survivors (Ricardo et al., 2011); ▶ Attitudes towards gender roles (Ricardo et al., 2011).

Men's skills	<ul style="list-style-type: none"> ▶ Skills that allow them to make changes to their behaviour as well as to address knowledge and attitudes; including tools to negotiate public face of masculinity and private face of masculinity (IPPF, 2010).
Men's behaviours	<ul style="list-style-type: none"> ▶ Objective measures (e.g. STI status for sexual and reproductive health programs; reports by partners about participants for healthy relationship programs; reports from women and other men about participants for violence programs) (IPPF, 2010); ▶ Self-reporting of men using a condom, those seeking medical treatment, reporting of symptoms over time (IPPF, 2010); ▶ Perpetration of rape or other forms of sexual violence against a girl or woman (Ricardo et al., 2011); ▶ Perpetration of non-sexual forms of violence against a girl or woman ▶ Bystander behaviors (Ricardo et al., 2011).

APPENDIX 2:

SAMPLE OUTCOME FRAMEWORK (MINERSON ET AL., 2011)

LEVEL OF OUTCOME	DESCRIPTION
Strengthening Individual Knowledge and Skills.	Education, skills building and awareness raising at an individual level.
Promoting Community Education.	Group efforts, social marketing and communications, media strategies.
Educating Professionals and Service Providers.	Training teachers, police officers, coaches, or doctors for example to do primary prevention work in their specific target communities.
Engaging and Mobilizing Communities.	Building coalitions and networks, identifying and building capacity of male leaders, awareness events, White Ribbon Campaigns.
Changing Organizational Practices.	Challenging and changing entrenched practices that tolerate or provide impunity for gender-based violence.
Influencing Policy and Legislation.	Legal and policy reform

APPENDIX 3:

SAMPLE INDICATORS	
Youth (Bloom, 2008)	<ul style="list-style-type: none"> ▶ Proportion of people who report they heard or saw a mass media message on issues related to sexual violence and youth ▶ Proportion of girls who say they would be willing to report any experience of unwanted sexual activity ▶ Proportion of girls that feel able to say no to sexual activity ▶ Proportion of girls reporting that male teachers do not have the right to demand sex from school children
Community mobilization and individual behavior	<ul style="list-style-type: none"> ▶ Proportion of people who have been exposed to VAW/G prevention messages, proportion of individuals who know any of the legal sanctions for VAW/G ▶ Know legal rights of women ▶ People who say that wife beating is an acceptable way for husbands to discipline their wives ▶ Proportion of people who agree that a woman has a right to refuse sex ▶ Proportion of people who agree that rape can take place between a man and woman who are married
Working with men and boys	<ul style="list-style-type: none"> ▶ Number of programs implemented for men and boys that include examining gender and culture norms related to GBV ▶ Proportion of men and boys who agree that women should have same rights as men ▶ Proportion of men and boys with gender-related norms that put women and girls at risk for physical and sexual violence ▶ Proportion of men and boys who believe that men can prevent physical and sexual violence against women and girls

Sample areas for process indicators (IPPF, 2010)

- ▶ Program sustainability;
- ▶ Cultural appropriateness of intervention methods;
- ▶ Program staff training - skill and confidence level in working with men; capacity of staff (for e.g. to deliver training, to assess interest and commitment among trainers);
- ▶ Meaningfulness and usefulness of intervention goals and outcomes to men;
- ▶ Integration of programming into existing structures and systems (e.g. subjects in school curricula, teaching and professional training);
- ▶ Positive message (correcting the 'faults' of men does not appeal to them);
- ▶ Monitoring of counselling quality; sufficient length of training relative to the material covered for trainers;
- ▶ Creation of a safe and reflective space for staff and participants to examine and challenge their own attitudes and mindsets as men and women.



white ribbon