We Can Bangladesh Effectiveness Review
Full Technical Report

Oxfam GB
Policy Influencing Outcome Indicator

March, 2012

Photo: G. M. B. Akash

Table of Contents

Executive Summary

As per Oxfam Great Britain's (OGB) Global Performance Framework (GPF), samples of mature projects are being randomly selected each year and their effectiveness rigorously assessed. A project focusing on the provision of support to the We Can Secretariat of Bangladesh was selected in this way under the policy influencing thematic area. However, rather than simply assessing the effectiveness of this support in terms of strengthening the work of the Secretariat, it was decided that it would be of greater interest to assess the effectiveness of the core campaign it was set up to spearhead. This campaign – the We Can Campaign – seeks to change deeply ingrained attitudes and practices that endorse gender discrimination and violence. It was initially launched in 2004 in six South Asian countries and has since been replicated in several African and Middle Eastern countries.

Having been implemented in many areas of Bangladesh, it was unrealistic to assess the effectiveness of the entire campaign. A decision was consequently taken to focus the impact assessment on the work of one of the We Can Alliance partners – *Polli Sree*. This organisation, based in Dijnapur district in Bangladesh's Rangpur division, was chosen particularly due to the high prevalence of domestic violence that takes place in its operational area.

In September 2011, with the support of an external consultant, specially designed questionnaires relating to intra-marital violence were administered to randomly selected samples of 1,159 women and 1,154 men in 92 villages situated in Thakugaon, Panchagar and Niphamari districts of Rangpur division and Naogaon district of Rajshahi division. Campaign activities were implemented in 42 of these villages, while the other 50 villages were selected for comparison purposes. Both questionnaires and the data collection process were informed by the World Health Organisation (WHO) and PATH's guidelines for researching violence against women. Statistical analysis was undertaken using propensity score matching (PSM) and multivariable regression (MVR) to control for measured differences between the women and men of the intervention and comparison villages.

Overall, statistically significant and positive differences were found between women and men residing in the implemented and non-implemented sites in relation to both gender and intra-marital violence attitudes (p-value < 0.001). However, when the data are disaggregated by research site, it is clear that these differences only apply to one site in particular. In this site the campaign was more intensely implemented. The lack of evidence of impact in the other sites, then, appears to be due to differences in implementation, rather than the We Can Campaign model per se. The other interesting – but perhaps not surprising – finding is that there is evidence that the campaign affected the attitudes of change makers to a greater extent than non-change makers in the site where it was more intensely implemented. Finally, in this particular site, women were more likely to report cessation of intra-marital violence (p-value < 0.05).

It is important to point out that the review's findings do not imply that the campaign had no impact on either the change makers and/or those in their immediate circles of influence in those sites where the implementation of the campaign was less intense. It is simply that there is no evidence that the campaign had an impact on the general population in these sites.

To strengthen the campaign's impact, the We Can Alliance is encouraged to consider the following:

- Identify key reasons for differences in campaign implementation in the Naogaon site vis-à-vis the Thakugaon/Panchagar and Nilphamari sites
- Explore ways to ensure that the We Can Campaign is carried out with significant intensity at the local level in the future, possibly sacrificing geographic scale
- Consider carrying out complementary qualitative research to interrogate, and possibly, challenge the We Can Campaign's theory of change
- Review either the design and/or implementation of the We Can "conscientisation" tools/processes for change makers
- Investigate possibilities further testing the effectiveness of the We Can Campaign model

1.0 Introduction and Purpose

Oxfam GB has put in place a Global Performance Framework (GPF) as part of its effort to better understand and communicate its effectiveness, as well as enhance learning across the organisation. This framework requires programme teams to annually report output data across six thematic indicator areas. In addition, modest samples of sufficiently mature projects (e.g. those closing during a given financial year) associated with each thematic indicator area are being randomly selected each year and rigorously evaluated. One key focus is on the extent they have promoted change in relation to relevant OGB global outcome indicators.

The global outcome indicator for the policy influencing thematic area is based on contribution scores generated from the findings of rigorous qualitative evaluations. This indicator is explained further in Section 3.0 below, and the work that took place in Bangladesh in September 2011 was part of an effort to capture data on this indicator. The original project randomly selected for the effectiveness review was entitled "We Can Campaign in Bangladesh: Alliance & Secretariat Management" (P00115). This specific project was set up to provide the We Can Campaign Secretariat in Bangladesh with financial support to cover management operations, communication materials development, and the organisation of events through allies and partners.

The Effectiveness Review of Bangladesh's We Can Campaign focused on the work spearheaded by OGB's partner, Polli Sree, in the north-western part of the country.

Given that an effectiveness review is, for all intent and purposes, a type of impact assessment, it made sense to focus the review on the effectiveness of the We Can Campaign itself. However, the work of the campaign in Bangladesh is not primarily focused on changing government policy but on changing popularly held patriarchal values, attitudes, and practices that perpetuate violence against women (VAW). The evaluation design that was adopted, therefore, sought to compare comparable geographic areas where campaign activities had and had not been implemented.

However, it was impractical to carry out the evaluation in all areas of the country where campaign activities have taken place. A decision was therefore made to focus on the work of a local partner organisation in one particular area of the country. The name of this local partner is *Polli Sree* (meaning "real beauty" in English) whose headquarters is in Dijnapur District, Rangpur Division in north-west Bangladesh. It has been spearheading the campaign in this part of the country since 2005, starting first with its home district of Dijnapur and then reaching out to several other surrounding districts.

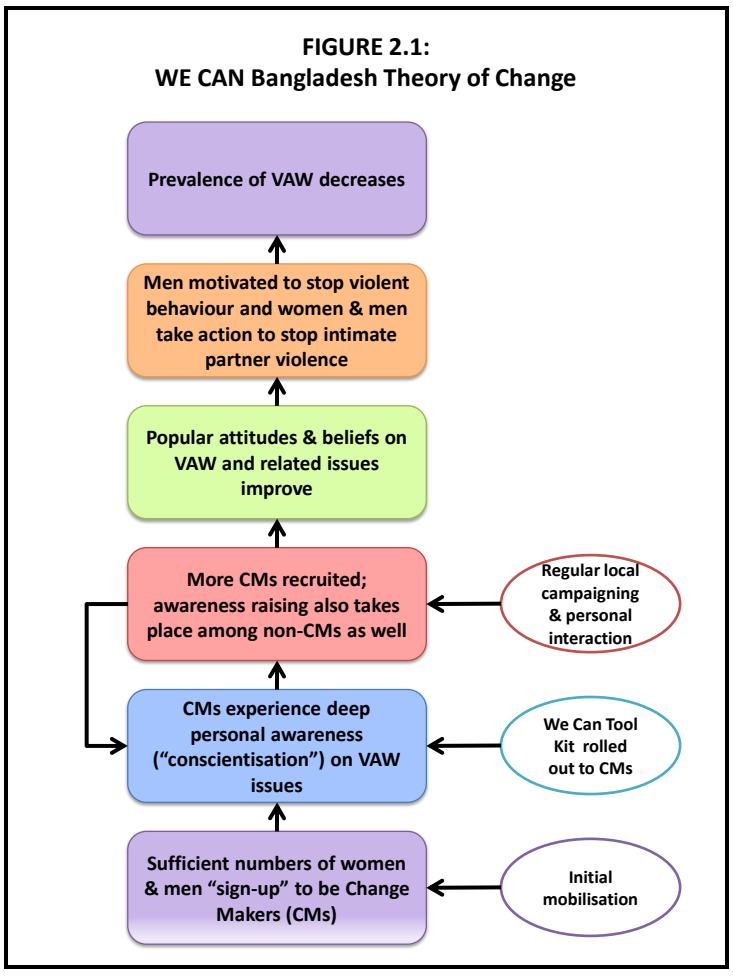
This report presents the findings emerging from a process where specially designed questionnaires were administered to men and women in areas where the campaign had been implemented and similar areas where it had not. However, before doing so, Section 3.0 explains the Bangladesh variation of the We Can Campaign's theory of change. Section 4.0, Section 5.0, and Section 6.0 follow by presenting the conceptual framework underlying the indicator, the impact evaluation design that was used, and the methods of data collection and analysis, respectively. Section 7.0 is the longest section of this document. Its subsections include those related to basic descriptive statistics, intervention exposure, and finally the overall differences between the intervention and comparison sites. Section 8.0 concludes the document with general conclusions and programme learning considerations.

2.0 Intervention Logic of the We Can Bangladesh Campaign

As part of the formative work that was undertaken to prepare for the data collection exercise, efforts were undertaken with both We Can Campaign Secretariat and *Polli Sree* staff to understand and unpack the intervention logic or "theory of change" underpinning the Bangladesh variation of the We Can Campaign. This is particularly relating to the work taking place at the grassroots level. The resulting theory of change emerging from this process is visually depicted in Figure 2.1.

The Bangladesh version of the campaign is unique in that change makers are continuously re-engaged and supported to campaign.

FIGURE 2.1:
WE CAN Bangladesh Theory of Change



In new areas (e.g. villages and wards), initial work is undertaken to recruit change makers. These are both women and men that purportedly have both desire and interest in tackling VAW issues in their communities. The recruited change makers then undergo training and several other reflection and awareness raising processes. They are encouraged to reflect on and change their own behaviour and then to encourage others to do the same. *"Each change maker commits to never tolerating or perpetrating violence against women in their own lives and to reach out to ten others in an attempt to influence their attitudes and practices regarding gender discrimination and gender-based violence."*[1]

After the change makers undergo this initial "conscientisation" process, they

[1] We Can *Campaign Strategy Paper, Updated March 2007*, page 7.

form committees at Ward level and are supported by local partner organisations like *Polli Sree* to carry out VAW campaigns in their communities. Through their campaigning and personal interaction with others, two things are intended to happen: First, more women and men are to sign up as change makers and, thereby, also experience deep rooted personal transformation. However, levels of awareness even among those who do not become change makers also increases, and they too come to recognise that VAW is unacceptable. Popular beliefs and attitudes are, consequently, improved. Men are thus motivated to change their behaviour, and both women and men take action to stop VAW. The final result is a decrease in the prevalence of VAW.

3.0 The Policy Influencing Outcome Indicator and Outcomes of Interest

3.1 The Policy Influencing Outcome Indicator

The methodology that is generally being used in the effectiveness reviews of OGB supported policy influencing work is informed by a qualitative research methodology know as *process tracing*. This approach is particularly suitable when the number of units an intervention is attempting to affect is small (e.g. a campaign that is attempting to change government policy). Here, efforts are first undertaken to assess the extent both the intended and unintended outcomes related to the intervention came about. The next step is to then evidence what factors (mechanisms) brought about these observed outcome changes. These factors may or may not be related to the intervention in question.

The extent there is evidence linking the campaign work to the observed outcome changes is then summarised by the external researcher as "contribution scores". These scores reflect two things: The first is the extent the outcome change targeted by the campaign actually came about. The second is the extent there is evidence that the campaign was responsible. If, for example, significant targeted outcome change is observed and there is clear evidence that this change was largely due to the workings of the campaign, a high score is given. On the other hand, a low score is assigned even if such change is observed but there is little evidence that the campaign was responsible. Middle-range scores are allocated if there is evidence that the campaign partly contributed to observed, targeted change and/or the targeted change only partially manifested. The "scoring key" that is used is presented in Annex 1.

OGB's main methodology for assessing the effectiveness of campaigning work was not used, given that the campaign is targeting large numbers of people.

However, the process tracing methodology can be difficult to apply when the number of units an intervention is attempting to affect is large, as is the case with many community-based interventions that are attempting promote positive change for large numbers of people. Such interventions are referred to as large *n* interventions in the literature.[1]

It is possible to practically apply a methodology such as process tracing to a small number people. One could explore whether, for example, a given household's income has increased, and, if so, employ rigorous qualitative

research methods to evidence how the increase came about. However, people interface with and are affected by the interventions of most social programmes in different ways. It is, therefore, typically not advisable to generalise the findings generated by such in-depth case studies to other programme participants. While it is theoretically possible carry out such intensive qualitative work on a representative sample of programme participants, feasibility considerations are likely to make this impractical.

3.2 Adapting the Indicator in the Context of the We Can Campaign in Bangladesh

The We Can campaign work that is taking place at the grass-roots level in Bangladesh is a large *n* intervention. As such and following the above, employing a qualitative impact assessment strategy such as process tracing would be difficult to successfully implement. Approaches typically used in the evaluation of large *n* interventions are thus more appropriate. The particular approach that was used is explained in the next section.

The impact assessment design was based on a framework used when the number of people being targeted is large.

What does this mean for the global policy influencing outcome indicator? Fortunately, the results generated from a large *n* impact assessment design can still be summarised as "contribution scores". The process is actually simpler because the two steps – evidencing the extent the targeted outcomes have come about and evidencing the extent the campaign was responsible – are merged. As will become clearer in the next section, the way the impact of a large *n* intervention is assessed is by comparing the intervention group with a control or comparison group. The control or comparison group is intended to represent what would have happened to the average person in the intervention group had they never been exposed to the intervention. Hence, the average difference between the two groups is intended to capture the "net change" the intervention group experienced as a result of the intervention.

3.3 Outcomes of Focus and their Measures

Drawing from the theory of change presented in Section 2.0, the impact of the We Can Campaign – as implemented in *Polli Sree's* working area – on three key outcomes was assessed:

- Gender attitudes
- Specific attitudes related to VAW
- Women's experience of different forms of intra-marital violence

What follows is a description of the various measures used for each of these outcomes:

- *Gender attitudes*

The respondents were asked to state their level of agreement – ranging from strongly agree to strongly disagree – to 10 different statements related gender roles. These statements included:

1. A wife should obey her husband, even if she disagrees with him.

2. Men should help with work around the house such as doing dishes, cleaning, looking after children, and so forth.
3. It is important for a husband to show his wife that he is the boss.
4. A wife should be able to choose her own friends even if her husband disagrees.
5. A man's job is to earn money; a women's job is to look after the home and family.
6. It's a wife's obligation to have sex with her husband even if she doesn't want to.
7. If a man mistreats his wife, others outside the family should intervene.
8. Women are as important as men in ensuring that the basic needs of families are met.
9. Men have a responsibility in childcare.
10. A man who works outside the home should not be expected to help with housework.

The campaign's impact on three specific outcome areas was assessed – gender attitudes, VAW attitudes, and women's experience of intra-marital violence .

As is apparent, some of these statements are positive from a gender perspective, while others are negative. If a respondent strongly agreed with a positive statement, for example, they would obtain the highest possible score. And the more they disagreed, the lower their score would be. The reverse was the case for negative statements, i.e. the greater the disagreement, the higher the score.

Rather than simply using the raw scores as the bases of the gender attitudes measure, principal factor analysis was carried out on the 10 items to generate factor scores. This approach narrows in on the variation in the data that is common in the responses. In so doing, it reduces the amount of "noise" present in the data, thereby, enhancing measurement precision.

- *Specific attitudes in relation to VAW*

Two modules were used to measure the respondents' VAW attitudes. The first was similar in structure as the gender attitudes measure but, of course, differed in its content. It comprised of the following statements in particular:

1. A husband has the right to hit his wife when she is disobedient.
2. There is no excuse for a man hitting a woman.
3. Some wives try to get beaten by their husbands in order to get sympathy from others.
4. A wife should move out of the house if her husband hits her often.
5. Government agencies and NGOs should do more to stop husbands from hitting their wives.
6. Even when wives lie to their husbands, they should not get beaten.
7. It does some wives some good to be occasionally hit by their husbands.
8. A man should be arrested if he hits his wife.
9. A husband has no right to hit his wife even if she breaks agreements she made with him.
10. Sometimes it is justifiable for a man to beat his wife.
11. Women should be protected by law if their husbands beat them.
12. Cases of wife beating are the fault of the husband, not the wife.

The other module was adapted from the World Health Organisation's (WHO) Multi-country Study on Women's Health and Domestic Violence against Women.[2] Here, the respondents were asked the following: *In your opinion, does a husband have good reason to hit his wife if*:

1. She spends money on things he does not approve of.
2. She goes outside of the home without his permission.
3. She talks back to him.
4. She disobeys him.
5. She refuses to have sex with him.
6. He suspects that she has been unfaithful.
7. She does not serve him as he expects to be served.

If the respondent answered yes to any of the above questions, s/he was coded as condoning VAW.

- *Women's experience of different forms of violence*

As will be explained in more detail in Section 5.0 below, female enumerators also asked female married respondents the extent they experienced intra-marital violence – both mental and physical – during the last 12 months. The particular questions asked were again adapted from WHO/PATH's multi-country study. The respondent was first asked whether her husband had subjected her to one of the items outlined below and, if yes, then a follow-up question was asked on the number of times, i.e. 1-2 times, 3-5 times, 6-10 times, or over 10 times. In particular, the women were asked: *Has your husband ever done any of the following in the past 12 months, that is, since the ending of Ramadan one year ago up to the present*:

The female respondents were asked if they had been subjected to specific actions that represent different types and degrees of intra-marital VAW .

1. Ignored you?
2. Tried to keep you from seeing your friends/neighbours?
3. Expected you to ask his permission before seeking healthcare for yourself?
4. Refused to give you enough money for household expenses, even when he had enough money?
5. Insulted you or made you feel bad about yourself?
6. Humiliated you in front of other people like other family members, neighbours, or others in the community?
7. Verbally threatened to hurt you or someone you care about?
8. Slapped you or thrown something at you that could hurt you?
9. Pushed you or pulled your hair?
10. Hit you with his fist or with anything else that could hurt you?
11. Kicked, dragged, or choked your or beat you up?
12. Attempted to do serious harm to any part of your body?
13. Actually inflicted serious damage to one or more parts of your body?
14. Physically forced you to have sexual intercourse even though you did not want to?

The specific questions were further subdivided into specific categories of abuse – mental abuse, general physical abuse, and serious physical abuse.

[2] http://www.who.int/gender/violence/who_multicountry_study/en/

Several of the introductory questions were asked to gradually lead the respondent to the more clear and pertinent questions. As such, their responses to these introductory questions did not inform any of the categories. Only the responses starting from question 5 were used. In particular, any positive response to either question 5 or 6 was coded as subjection to mental abuse, While any positive response to questions 7 through to 14 as subjection to physical abuse. Finally, a positive response to any of the questions from 10 to 14 was coded as subjection to serious physical abuse.

To control for baseline differences related to VAW experience, efforts were undertaken to reconstruct baseline data.

The women who had been married for more than six years, i.e. since before the commencement of the We Can Campaign, were re-asked these same questions, but this time with the 12 month period associated with 2004 being the reference period. Historical reference points were used to help the women recall back to this particular year. This was done in an attempt to reconstruct baseline data on women's experience of intra-marital violence.

4.0 Impact Assessment Design

4.1 Limitations in Pursuing the Gold Standard

The core challenge of large *n* social impact evaluations is to credibly estimate the net effect of an intervention or programme on its participants. An intervention's net effect is typically defined as the average gain participants realise in outcome (e.g. reduced subjection to violence) from their participation. In other words:

Impact = average post-programme outcome of participants – what the average post-programme outcome of these same participants would have been had they never participated

This formula seems straightforward enough. However, *directly* obtaining data on the latter part of the equation – commonly referred to as the counterfactual – is logically impossible. This is because a person, household, community, etc. cannot *simultaneously* both participate and not participate in a programme. The counterfactual state can therefore never be observed directly; it can only be estimated.

The randomised experiment is regarded by many as the most credible way of estimating the counterfactual, particularly when the number of units (e.g. people, households, or, in some cases, communities) being targeted is large. The random assignment of a sufficiently large number of such units to intervention and control groups should ensure that the statistical attributes of the two resulting groups are similar in terms of a) baseline outcome status (e.g. both groups have the same average incomes); and b) both their observed characteristics (e.g. education levels) and unobserved characteristics (e.g. motivation) that affect the outcome variables of interest. In other words, randomisation works to ensure that the *potential outcomes* of both groups are the same. As a result – provided that threats such differential attrition and intervention spill-over are minimal – any observed outcome differences

observed at follow-up between the groups can be attributed to the workings of the programme.

Unfortunately – outside the context of specially designed pilot studies – randomised evaluation designs are seldom implemented in the context of social programmes, particularly in low-income countries and in the non-governmental (NGO) sector. There can be cost, feasibility, and/or ethical constraints that militate against their use or simply the desire among implementing agencies to work with purposively chosen populations. Moreover, there are often cases where the *opportunity* to participate in a programme is put in place – as would be the case with the setting up of a micro-credit programme – and people *choose* whether to participate. Those who choose to participate are likely to be different than those who do not, including in characteristics that are intrinsically difficult to measure, e.g. motivation.

4.2 Alternative Evaluation Design Pursued

There are several evaluation designs when the comparison group is non-equivalent that can – particularly when certain assumptions are made – identify reasonably precise intervention effect estimates. One solution is offered by matching: Find units in an external comparison group that possess the same characteristics, e.g. ethnicity, age, and sex, as those of the intervention group and match them on these characteristics. If matching is done properly in this way, the observed characteristics of the matched comparison group will be identical to those of the intervention group.

The problem, however, with conventional matching methods is that with large numbers of characteristics on which to match, it is difficult to find comparators with similar combinations of characteristics for each of the units in the intervention group. The end result, typically, is that only a few units from the intervention and comparison groups get matched up. This not only significantly reduces the size of the sample but also limits the extent the findings can be generalised to all programme participants. (This is referred to as the "curse of dimensionality" in the literature.)

The evaluation design involved comparing areas where the campaign had been implemented and not implemented, while statistical procedures were used to control for potentially confounding factors.

Fortunately, matching on the basis of the propensity score – the conditional probability of being assigned to the programme group, given particular background variables or observable characteristics – offers a way out. The way propensity score matching (PSM) works is a follows: Units from both the intervention and comparison groups are pooled together. A statistical probability model is estimated, typically through logit or probit regression. This is used to estimate programme participation probabilities for all units in the pooled sample. Intervention and comparison units are then matched within certain ranges of their conditional probability scores. Tests are further carried out to assess whether the distributions of characteristics are similar in both groups after matching. If not, the matching bandwidth or calliper is repeatedly narrowed until the observed characteristics of the groups are statistically similar. Provided that a) the dataset in question is rich and of good quality; b) the groups possess many units with common characteristics (i.e. there is a large area of common support); and c) there are no unobserved

differences lurking among the groups, particularly those associated with the outcomes of interest, PSM can generate good intervention effect estimates.

Two popular methods were used to address selection bias – propensity score matching and multivariable regression.

Multivariable regression is another approach that is also used to control for measured differences between intervention and comparison groups. It operates differently from PSM in that it seeks to isolate the variation in the outcome variable explained by being in the intervention group *net of other explanatory variables* (key factors that explain variability in outcome) included in the model. In this way, multivariable regression controls for measured differences between the intervention and comparison group. The validity of both PSM and multivariable regression are founded heavily on the "selection on observables" assumption, and, therefore, treatment effect estimates can be biased if unmeasured (or improperly measured) but relevant differences exist between the groups.[3] Both PSM and multivariable regression were employed during data analysis, and efforts were made to capture key explanatory variables believed to be relevant in terms of the assessed outcomes, e.g. sex and age of household head, educations levels, outstanding dowry debts, etc. (see Section 6.0 below).

While no baseline data were available, efforts were made, as explained above, to reconstruct it through respondent recall. This method does have limitations, e.g. memory failure, confusion between time periods, etc. However, for data that can be sensibly recalled, e.g. ownership of particular household assets, recall methods can aid in enhancing the validity of a cross-sectional impact evaluation design. The reconstructed baseline data were used in two ways. First, several of the variables included in the PSM and regression procedures were baseline variables constructed from recalled baseline data. One set of variables, for example, was related to the respondents wealth status at baseline, e.g. whether they were asset rich, asset poor, or somewhere in between. This was done in attempt to control for baseline wealth differences between households in the intervention and comparison villages.

The second way the reconstructed baseline data were used was to derive pseudo double-difference intervention effect estimates. With longitudinal or panel data, this is implemented by subtracting each unit's baseline measure of outcome from its endline measure of outcome (i.e. endline outcome status minus baseline outcome status). The intention here is to control for time invariant differences between the groups. Bearing in mind the limitations associated recalled baseline data, using PSM and/or regression and the double-difference approaches together is considered a strong impact evaluation design.

4.2 Control Variables

Key to the success of the above evaluation design is being able to control for key differences between the respondents in the intervention and comparison

[3] One of the MVR procedures that was used attempted to control for possible unobserved differences between the groups. This is the Heckman Selection Model or 2-step Estimator. Here, efforts are made to directly control for the part of the error term associated with the participation equation that is correlated with both participation and non-participation. The effectiveness of this method, however, depends, in part, how well the drivers of participation are modelled.

sites, particularly those relevant to VAW. The particular control variables for which data were collected are presented in Table 6.1.1 and Table 6.1.2 below.

Aside from standard demographic variables that are often used in such studies, efforts were undertaken to review the literature to identify key determinants or predictors of VAW, particularly those relevant to the Bangladesh context. Key predictors cited in the literature include:

Efforts were made to capture data on and control for key factors cited in the literature that are predictors of VAW.

- Education levels of both partners
- Age of woman, with women in new marriages being at greater risk
- Poverty status of household
- Man's past exposure to family violence
- Women's autonomy over income, with such autonomy being positively associated with VAW experience
- Urban/rural residence
- Family structure, with women in nuclear household's more likely to be subjected to VAW than those living with husbands' extended families
- Religion
- Substance use of husband
- Women in dowry arrangements, particularly those that have not been fulfilled
- Women living in "conservative" versus "progressive" areas
- Level of spousal communication[2, 3]

Data were collected on all of these predictors, save for the last two. In particular, efforts were not made to classify villages by their levels of cultural and religious conservatism. This was because it was assumed that both the intervention and comparison areas within each of the study areas were similar in this respect. However, degree of remoteness of the village in question was deemed to be an important factor, with the assumption being that more remote villages are more conservative.

Efforts were furthermore not undertaken to collect data on the level of spousal communication. The main reason was because this could be one of the outcomes associated with the campaign's work. Consequently, controlling for it could mask the impact of the programme. One can assume that improvements in attitudes towards women could increase their status in the household, thereby, resulting in increases in spousal communication.

4.4 The Comparison Population

A key factor in ensuring the validity of any non-randomised, large *n* impact evaluation design is to use an appropriate comparison group. This is particularly true for ex-post, cross-sectional designs. Comparators who differ in relevant baseline characteristics and/or who are subjected to different external events and influences will likely result in misleading conclusions about programme impact. Identifying a plausible comparison group is therefore critically important and is, generally speaking, not an easy task in non-experimental work.

The challenge we confronted, then, was how to identify areas that could be comparable with those where *Polli Sree* implemented the campaign.

Consequently, considerable time was spent mapping out areas in the five operational districts were the campaign was implemented. Most of the areas in the district where *Polli Sree* is based – Dijnapur – had been saturated with campaign activities, while more fragmented implementation had taken place in the other four surrounding districts. Not wanting to compare people coming from different districts, a decision was taken to exclude Dijnapur from the study and concentrate only on these four districts.

In these districts, efforts were then taken to identify areas (union councils) where the campaign had been successfully implemented and match them with similar union councils that had not been reached.[4] A total of six intervention union councils were identified and each was matched with two comparison union councils. Data were collected from a total of 92 villages. Key criteria used in identifying the comparison union councils included proximity to the matched intervention council and similar proximity to the main district road.

The comparison population was identified by matching areas where the campaign had been successfully implemented with similar areas where it had not.

5.0 Methods of Data Collection and Analysis

5.1 Data Collection

Two questionnaires – one for married male respondents and one for married female respondents – were developed and translated into Bengali to capture data on both the outcome variables presented in Section 3.0 above. Data for other key characteristics of the interviewed women and men were also obtained to implement the evaluation design described in Section 4.0. The questionnaires were pre-tested by *Polli Sree* field staff and the Consultant and subsequently revised.

Given the sensitive nature of the women's questionnaire in particular, an attempt was made during the piloting process to use a special technique, the randomised response model (RRM), to elicit responses to questions about intra-marital violence experience.[5] The aim here was to protect the confidentiality of the female respondents and hopefully obtain more reliable data. However, the Consultant who witnessed the piloting activity – a professional anthropologist – observed that the respondent did not react favourably to the process. She suspected that the respondent believed that the people interviewing her were attempting to trick her into confessing that

[4] In Bangladesh's administrative set-up, the smallest to highest levels of administrative units are: Village, Ward, Union Council, Upazilla, District, and Division. Union councils, in particular, are typically comprised of 10-15 villages and represent the lowest unit of local government.

[5] The simplest example of RRM is when the desired information is binary (yes/no) in nature, e.g. whether the household owns cattle or not, rather than a specific number. To access the sensitive information, the respondent is presented with a "randomisation device" such as a coin. The respondent is then instructed to utilise the randomisation device (e.g. flip the coin) and keep the resulting outcome confidential. S/he is instructed to answer "yes" for a particular outcome (e.g. the coin lands heads up) – *regardless of the actual truth*. If the output of the device goes the other way (e.g. the coin lands tails up), the respondent is directed to answer the question truthfully. In this way, if the respondent answers "yes" to the question, the interviewer has no way of knowing whether s/he said "yes" because it is the truth or because of the outcome of the randomisation device. While there is no way of knowing the truth with respect to a particular respondent, the average response for all the respondents combined can be obtained using a simple mathematical formula.

her husband beats her. Any further attempts to use RRM were subsequently abandoned.

The 24 enumerators – 15 females and nine males – that administered the questionnaires were primary university students or recent university graduates. Through her university networks, the Consultant identified approximately one-third of them and they travelled in from Dhaka to the survey area. The remainder were recruited from within Dijnapur. Given that the questionnaire for men was shorter, fewer male enumerators were recruited. Furthermore, one of the older and more experienced enumerators was given the responsibility of supervising the data collection process in one of the more distant districts, Naogaon. Approximately, 28 prospective enumerators completed the three day training course, which was led by the Consultant but also support by OGB staff. The second day involved a practice run at administering the questionnaires, followed by critical review of the performance of the enumerators. Several of them were subsequently disengaged.

As mentioned above, the questionnaires were administered in four districts. These districts included Panchagarh, Thakurgaon, Nilphamari, and Naogaon. The location of these districts is presented in Figure 5.1.1. Given that only two intervention and four comparison union councils were to be surveyed in both Panchagarh and Thakurgaon districts, these districts were combined together to form one survey area. Niphamari and Naogaon districts were each treated as separate survey sites. Survey teams of eight enumerators (five female and three male) were each assigned to one of three survey areas.

To select interviewees in each of the 92 villages, a three-stage sampling technique was used. In the first stage, government village population statistics were used to identify the number of respondents to be interviewed in each village using the probability proportionate to size (PPS) method.[6] To identify the targeted number of female and male respondents in each village (the second stage), local informants first mapped out the settlements that existed in the villages, as well as the approximate numbers of households contained in each. PPS sampling was again used to identify sampling quotas for each settlement. The enumerators implemented the third stage of the sampling strategy when they reached their assigned settlements. Here, they randomly selected a household to start the first interview through the "spin-the-pen" technique. They then went on to interview the third next household, as per the pen's original direction of travel. They continued to do this until their sampling quotas for their assigned settlements were exhausted.

Sampling was done in three stages. The first two were based on the PPS method, While the latter involved simple random sampling .

The work of the enumerators was closely monitored and scrutinised. A survey team leader was appointed in each group to initially check all completed questionnaires. These were then checked again by the Consultant and, on the first day of the survey, by OGB staff.

[6] link to PPS document

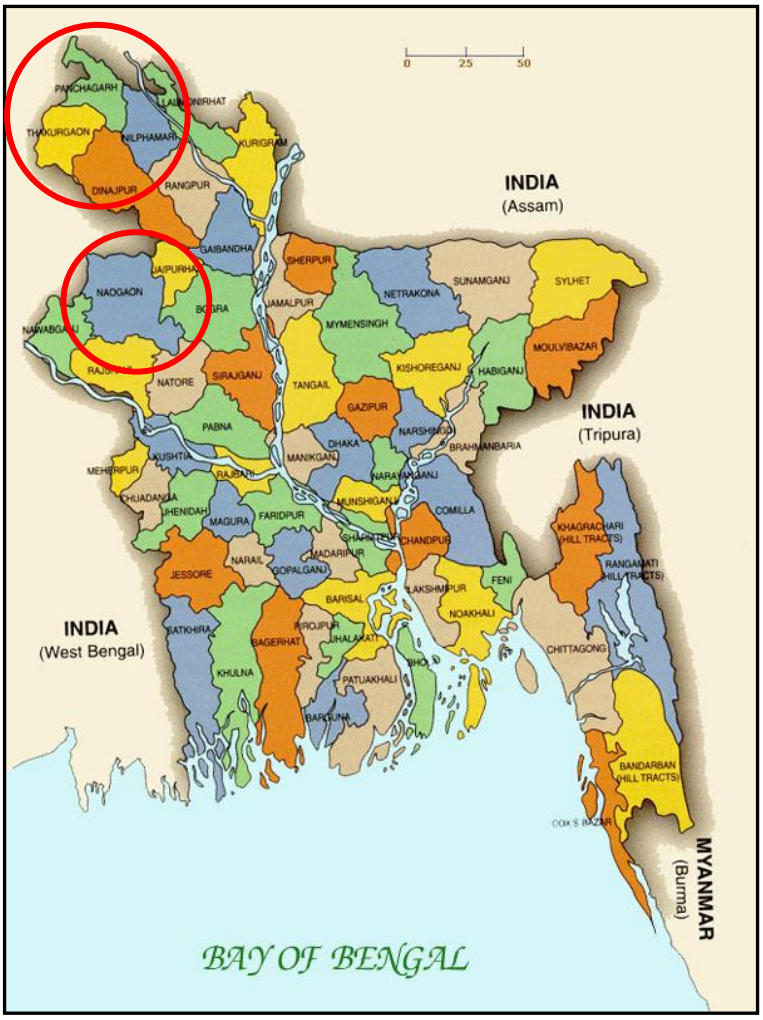


FIGURE 5.1.1: Location of We Can Survey Sites

5.2 Ethical Considerations

Efforts were undertaken to closely follow established ethical protocols for VAW research.

Given the sensitive nature of the subject matter, considerable efforts were undertaken to mitigate any potential negative impacts associated with the data collection process. In particular, the ethical protocols contained in WHO and PATH's publication entitled *Researching Violence Against Women: A Practical Guide for Researchers and Activists* were followed.[7] The specific steps taken included:

- Training the enumerators on research ethics in general and research related to VAW in particular, as informed by the above guide.
- Ensuring the informed voluntary consent was obtained before commencing the interviews.
- Carrying out the interviews in a private place, with females interviewing females and males interviewing males. (Training was provided to enumerators on making use of "dummy questions" in cases where the privacy was temporarily interrupted.)
- Only one respondent was interviewed from each household, particularly to avoid interviewing both the husband and the wife from the same household.

[7] http://www.who.int/reproductivehealth/publications/violence/9241546476/en/index.html

- The female interviews were concluded on a positive note to stress the respondent's strengths, and the respondents were provided with a list of service organisations in the local area that could offer support if required.

5.3 Data Analysis

OGB developed data entry tools in Adobe Acrobat Pro, and the Consultant recruited and supervised data entry clerks to enter the data. After identifying and rectifying some minor errors in MS Excel, the data were then imported into Stata for analysis, the results of which are presented in the following sections. Most of the analyses involved group mean comparisons using *t*-tests, as well as PSM with Stata's *psmatch2* module and various regression approaches.

Kernel and nearest neighbour matching without replacement were the main methods used in implementing PSM. Variables used in the matching process were identified by first using backwards stepwise regression to identify those variables that are correlated with the outcome measure of interest at *p*-values of 0.10 or less. The short-listed variables were then put into another stepwise regression model to identify those that are correlated with being a member of the intervention group. Covariate balance was checked following the implementation of each matching procedure. When covariate imbalance at *p*-values of 0.10 or less was identified, the bandwidth or calliper was reduced and the PSM procedure and covariate balance test implemented again. This was continued until all covariates were balanced at *p*-values greater than 0.20. Boot-strapped standard errors enabled the generation of confidence intervals to assess the statistical significance of the effect sizes. Exact matching within each survey area was further imposed to avoid comparing intervention and comparison respondents from different sites. Separate propensity scores were also generated for the female and male respondents of each district. This was done to ensure that the observable characteristics of the of both the male and female respondents were balanced, so that disaggregated PSM effect estimates by sex could be generated.

Data analysis was carried out centrally at OGB's head office using five different non-experimental estimation procedures.

All the covariates, as presented in Table 6.1.1 below, were included in the various regression approaches undertaken, i.e. regression with robust standard errors (to address issues of heteroskedasticity), robust regression (to reduce the influence of outliers), and regression with control functions (to attempt to control for relevant unobserved differences between the intervention and comparison men and women). To control for unobservable district and respondent sex specific influences, fix effect models were used, with the variables "district" and "respondent sex" specified as two key fixed effects.

5.4 Main Problems and Constraints Encountered

Overall, despite the usual hardships encountered when undertaking such intensive work, the data collection process went well. However, several challenges were encountered. These included:

- *Observable differences between the men and women of intervention and comparison villages*

Despite the efforts made to purposively match the intervention union councils to similar comparison union councils, some observable differences between the villages and respondents of the intervention and comparison groups were identified. While such observable differences are typically expected in non-experimental studies, they do have implications for data analysis and interpretation. This is elaborated upon further in Subsection 6.1 below.

- *Lack of campaign implementation intensity in three of the four districts*

The objective of the impact assessment was **not** to assess the overall impact of *Polli Sree's* work in relation to the campaign. Rather, it was to assess the effectiveness of the campaign in areas where it had been well implemented. In particular, efforts were made to work with *Polli Sree* staff to identify areas where the campaign had been well implemented and areas where it had not been implemented at all. In other words, the evaluation was focused on answering the question: "What happens when the We Can Campaign is well implemented?"

However, as is revealed below, there is little evidence that the campaign had been implemented with significant levels of intensity in two out of the three survey sites. This does, however, present a serendipitous opportunity to compare villages where the campaign had been intensely implemented with those where it had not.

6.0 Results

6.1 General Characteristics

Table 6.1.1 presents average statistics for general household characteristics obtained through the administration of the questionnaires to the respondents of both the intervention and comparison groups. The stars beside the number indicate differences between the two groups that are statistically significant at a 95 percent confidence level or greater. As is evident, while there are not many differences between the groups, there are some that are noteworthy. These include:

A number of observable differences were identified between respondents from the intervention and comparison sites .

- *Educational differences*.

Respondents from the intervention sites are slightly less likely to be uneducated and more likely to possess secondary education than their counterparts in the comparison sites.

- *Religious differences*.

Respondents from the intervention sites are less likely to be Muslim and more likely to be Hindu than those in the comparison sites.

- *Differences in group participation*.

Respondents from the intervention sites are more likely to be a member of a community group or some other organisation.

- *Wealth differences*.

Respondents from the Niphamari district intervention site in particular are more wealthy than those in the comparison sites.

- *Differences in recalled baseline violence experience in Niphamari district.*

No overall statistically significant differences in recalled, self-reported baseline violence experience were identified. However, significant differences were identified in Niphamari district, with fewer female respondents in the intervention sites reporting having been subjected to violence in the baseline period.

TABLE 6.1.1:
Descriptive Statistics: Intervention and Comparison Respondents Interviewed

			Overall		Thakurgaon/ Panchagar		Nilphamari		Naogaon	
	Intervention	Comparison								
	mean	**mean**	*dif.*	*t-stat.*	*dif.*	*t-stat.*	*dif.*	*t-stat.*	*dif.*	*t-stat.*
Respondent age	34.04	34.35	-0.31	-0.67	-0.36	-0.46	0.24	0.28	-0.65	-0.83
Respondent polygamous	0.01	0.01	0.00083	0.20	-0.0064	-1.36	0.0045	0.40	0.0035	0.60
Respondent uneducated	0.45	0.50	-0.042*	-1.97	-0.045	-1.20	-0.088*	-2.22	-0.0097	-0.28
Respondent primary school	0.32	0.31	0.013	0.65	-0.024	-0.67	0.060	1.69	0.012	0.36
Respondent secondary school	0.15	0.12	0.031*	2.11	0.057*	2.18	0.041	1.74	0.0018	0.07
Respondent post secondary	0.07	0.07	0.000022	0.00	0.014	0.75	-0.0085	-0.40	-0.0062	-0.34
Respondent Muslim	0.81	0.84	-0.034*	-2.12	0.13***	4.22	-0.092**	-3.14	-0.13***	-5.33
Respondent Hindu	0.19	0.15	0.036*	2.27	-0.12***	-4.16	0.098***	3.35	0.13***	5.41
Respondent other religion	0.00	0.00	-0.0021	-0.81	-0.0021	-0.79	-0.0051	-1.17	0.00047	0.09
Respondent ethnic minority	0.03	0.02	0.011	1.73	-0.0064	-1.36	0.0037	1.21	0.033*	2.13
Respondent good health	0.99	0.99	-0.0017	-0.36	-0.0018	-0.26	-0.0084	-0.91	0.0040	0.53
Age of HH head	42.53	42.65	-0.12	-0.22	-2.90**	-3.17	2.68*	2.48	0.16	0.17
Elderly headed household	0.00	0.00	-0.00072	-0.80	0	.	0	.	-0.0019	-0.78
Head has secondary education	0.18	0.18	0.00036	0.02	0.054	1.83	0.0017	0.06	-0.047	-1.72
# of productive adults	2.77	2.79	-0.015	-0.30	-0.24**	-2.80	0.25*	2.53	-0.018	-0.23
Household size	7.12	7.12	0.0014	0.01	-0.35	-1.90	0.46*	2.11	-0.070	-0.38
Number of adults	2.79	2.80	-0.015	-0.30	-0.25**	-2.94	0.26**	2.65	-0.017	-0.21
Number of children	4.34	4.32	0.017	0.20	-0.099	-0.75	0.20	1.33	-0.054	-0.38
Nuclear headed household	0.71	0.70	0.011	0.55	0.10**	2.89	-0.053	-1.52	-0.023	-0.72
Spouse of respondent is head	0.43	0.42	0.016	0.74	0.032	0.87	-0.0012	-0.03	0.014	0.40
Spouse age difference	-0.16	0.01	-0.17	-0.41	0.51	0.70	-0.84	-1.06	-0.26	-0.41
HH farms	0.75	0.75	0.0023	0.12	0.054*	2.00	-0.016	-0.40	-0.019	-0.64
HH rears livestock	0.86	0.85	0.0046	0.30	0.0086	0.37	0.036	1.14	-0.020	-0.79
HH hunts or fishes	0.14	0.14	0.0077	0.51	-0.020	-0.89	0.047	1.53	-0.0020	-0.08
HH runs business	0.35	0.33	0.023	1.14	0.029	0.84	0.031	0.85	0.016	0.46
HH does casual labour	0.32	0.35	-0.029	-1.43	-0.0017	-0.05	-0.073	-1.89	-0.023	-0.70
HH does unskilled wage labour	0.13	0.12	0.011	0.75	-0.015	-0.61	0.060*	2.14	-0.0078	-0.35
HH part of savings/credit group	0.43	0.42	0.012	0.58	-0.064	-1.75	0.099*	2.56	0.013	0.37
HH does skilled wage labour	0.10	0.10	0.0028	0.22	-0.0017	-0.07	0.041	1.74	-0.023	-1.16
Res. does domestic work	0.57	0.55	0.020	0.94	0.068	1.83	0.075	1.90	-0.064	-1.84
Res. farms	0.51	0.50	0.0064	0.30	0.059	1.61	-0.022	-0.55	-0.013	-0.37
Res. rears livestock	0.67	0.69	-0.020	-1.02	-0.00078	-0.02	-0.028	-0.72	-0.026	-0.84
Res. hunts	0.07	0.08	-0.0084	-0.75	-0.024	-1.55	-0.025	-1.00	0.015	0.86
Res. runs business	0.21	0.18	0.029	1.71	0.038	1.44	0.0054	0.18	0.042	1.38
Res. does casual labour	0.16	0.16	-0.0028	-0.18	0.010	0.37	-0.029	-0.96	0.0048	0.19
Res. does skilled labour	0.04	0.04	0.0056	0.66	0.0039	0.28	-0.010	-0.61	0.019	1.41
Res. part of savings/credit group	0.29	0.27	0.017	0.87	-0.033	-1.02	0.091**	2.62	0.0043	0.13
Res. does unskilled labour	0.05	0.04	0.0072	0.83	-0.0086	-0.65	0.024	1.44	0.0080	0.52
Res. in general group	0.09	0.04	0.053***	5.40	-0.0038	-0.52	0.076***	3.50	0.084***	4.49
HH owns land 2011	0.61	0.56	0.050*	2.36	0.099**	2.82	0.082*	2.07	-0.014	-0.39
Asset index 2011	0.20	-0.13	0.33**	2.67	0.15	0.75	0.94***	3.84	0.00096	0.00
Asset index 2004	0.15	-0.09	0.24*	2.11	0.16	0.83	0.82***	3.65	-0.14	-0.78
Marriage registered	0.69	0.72	-0.035	-1.76	0.084*	2.37	-0.10**	-3.06	-0.092**	-2.87
Dowry marriage	0.76	0.74	0.024	1.28	0.038	1.23	-0.067*	-2.56	0.073*	2.12
Dowry not paid	0.36	0.38	-0.028	-1.33	-0.061	-1.78	0.056	1.51	-0.060	-1.70
Number of year in marriage	12.68	12.95	-0.27	-0.59	0.020	0.03	-0.23	-0.30	-0.52	-0.64
Substance use of husband regular	0.06	0.08	-0.018	-1.12	-0.056*	-2.02	0.025	1.22	-0.014	-0.48
Violence res. father on mother	0.33	0.36	-0.034	-1.17	-0.015	-0.30	-0.15**	-2.88	0.043	0.93
Violence hus.' father on mother	0.29	0.26	0.026	0.79	0.092	1.56	-0.041	-0.78	0.034	0.59
Physical violence baseline	0.69	0.71	-0.023	-0.67	0.0083	0.16	-0.24***	-4.27	0.11	1.93
Serious violence baseline	0.50	0.55	-0.052	-1.42	-0.034	-0.56	-0.18**	-2.78	0.025	0.44
Serious violence no forced sex	0.21	0.25	-0.044	-1.41	0.029	0.50	-0.15**	-2.60	-0.020	-0.42
Observations	**883**	**1388**	**2271**		**757**		**656**		**858**	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.1.2 presents population and location related statistics derived from the data obtained through the compilation of village data capture forms completed by the survey team leaders. As is apparent from the table, the intervention villages are, on average, smaller in terms of both population and size than are the comparison villages. They are also less remote, given that they are closer to the nearest municipality and main district road.

TABLE 6.1.2:
Descriptive Statistics: Intervention and Comparison Villages Surveyed

			Overall		Thakurgaon/ Panchagar		Nilphamari		Naogaon	
	Intervention mean	Comparison mean	*dif.*	*t-stat.*	*dif.*	*t-stat.*	*dif.*	*t-stat.*	*dif.*	*t-stat.*
Population of village	2293.55	3367.96	-1074.4*	-2.52	-718.9	-1.11	-2425.6**	-3.40	120.4	0.15
Number of HHs in village	483.48	595.48	-112.0	-1.17	-200.6	-1.02	-359.2*	-2.58	228.9	1.73
Village area	2.36	5.82	-3.46	-1.56	-9.85	-1.47	-1.45	-1.78	-0.83	-1.63
Distance nearest municipality	7.25	16.15	-8.90***	-5.57	-12.5**	-3.20	-8.95***	-6.07	-5.92*	-2.25
Distance from district road	3.11	5.45	-2.34*	-2.33	-2.54	-1.25	-5.10***	-4.17	1.60	0.82
Distance from union centre	2.96	2.38	0.58	1.20	0.23	0.41	0.47	0.92	1.63	1.34
Distance from district centre	17.06	21.36	-4.30	-1.54	-10.6*	-2.61	-10.2*	-2.35	11.5*	2.24
Observations	**42**	**50**	**92**		**29**		**35**		**28**	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Given that we are interested in estimating the impacts of the We Can Campaign in the study areas, the above differences between the intervention and comparison sites are a cause for concern. This is particularly because many of them – as presented in Subsection 4.2 above – are documented in the literature as being predictors of intimate partner violence. The particularly noteworthy differences include those related to:

- Educational levels of the respondents
- Religion
- Household wealth
- Baseline violence experience, particularly for Nilphamari district

It is worth noting, however, that many of the other determinants of VAW cited in the literature, e.g. dowry not paid, are statistically balanced between the intervention and comparison groups.

Given that there are relevant differences between the intervention and comparison groups, directly comparing them may very well result in biased estimations of the impacts of the We Can Campaign. Consequently, it was critical to control for these differences in the statistical analysis of the data.

6.2 Campaign Implementation

Data were collected pertaining to campaign exposure in two ways. First, a village data capture form was used to obtain data from village informants on a number of issues pertaining to the campaign's implementation. These include:

Efforts were also made capture and analyse data on the extent the campaign was implemented in the three survey sites.

- Whether change makers in either the village or anywhere in the Ward where the village is located were present, and whether they were organised into committees.
- The numbers of female and male change makers present in the village and ward, as well as the number of years they have been active.
- The number of campaigns the change makers carried out at the village and ward levels, including the number of people reached by these campaigns.

The results are summarised in Table 6.2.2 below. According to the village level informants, approximately 60 percent of the surveyed intervention villages and Wards were reported to have change makers present. However, there are considerable differences among the survey sites. In Thakurgaon/

Panchagar and Nilphamari less than half of the intervention villages reported having change makers. However, they were reported to be present in *all* the villages in the Naogaon site. The average number of change makers reported to be present at the Ward and village levels are further reflective of this trend.

TABLE 6.2.2:
Statistics on Intensity of We Can Campaign at Ward and Village Levels

	Overall	Thakurgaon/ Panchagar	Nilphamari	Naogaon
Proportion of Wards with CMs	0.57	0.40	0.44	1.00
Proportion villages with CMs	0.60	0.47	0.44	1.00
Number years CM in Ward	1.44	0.31	1.50	2.91
Average number of CMs in Ward	33.88	3.27	14.06	104.45
Avg. # of female CMs in Ward	19.43	1.87	6.56	62.09
Avg. # of male CMs in Ward	14.45	1.40	7.50	42.36
Average # of CMs in village	35.88	2.93	5.88	124.45
Avg. # of female CMs in village	21.10	1.73	2.63	74.36
Avg. # of male CMs in village	14.79	1.20	3.25	50.09
CM campaign in Ward	0.36	0.00	0.31	0.91
Avg. # CM campaign in Ward	0.93	0.00	1.00	2.09
CM campaign carried out in vil.	0.33	0.00	0.31	0.82
Average # of vil. campaigns	0.79	0.00	1.00	1.55
# of villages	**42**	**15**	**16**	**11**

In cases where the village informants reported the presence of change makers, either at the Ward or village level, they were asked to report the extent to which they had been campaigning. As indicated in the table, campaigns were reported to have taken place in only about one-third of the villages and Wards. Again, there are significant differences between the districts. No campaigns, for instance, were reported to have taken place in the Thakurgaon/ Panchagar site. This is contrasted with an average of one and two campaigns in the Nilphamari and Naogaon sites, respectively.

The data clearly reveal that the campaign was more intensely implemented in the Naogaon site.

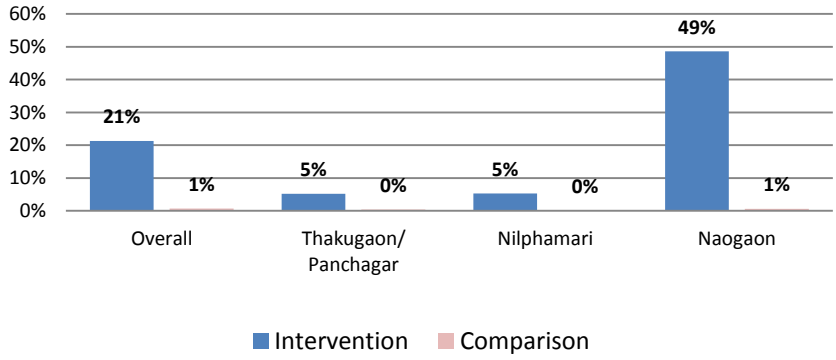
At least according to the village informants, it appears that the campaign had been more active in the Naogaon site than in the other sites. However, it is possible that the survey teams – by chance – interviewed informants from the Thakurgaon/Panchagar and Nilphamari sites that were not fully aware of the campaign's activities.

We can triangulate the data provided by the village informants in several different ways. In the questionnaires, for example, the respondents were asked at the end of the interview whether they and/or their spouse is a change maker *with an identification card*.[8] Given that the respondents were randomly selected in the sites, we would therefore expect a higher proportion of the respondents to have reported themselves or their spouses as being change makers in the Naogaon site, if there were really more change makers in this site. Figure 6.2.1 presents the relevant statistics. As indicated, only five percent of the respondents in the first two sites reported that they and/or their spouses were change makers, compared with nearly half in Naogaon.

[8] In Bangladesh, it is not simply a matter of one identifying oneself as a change maker or not. The process is more formalised, and those successfully indoctrinated as change makers are issued with identification cards.

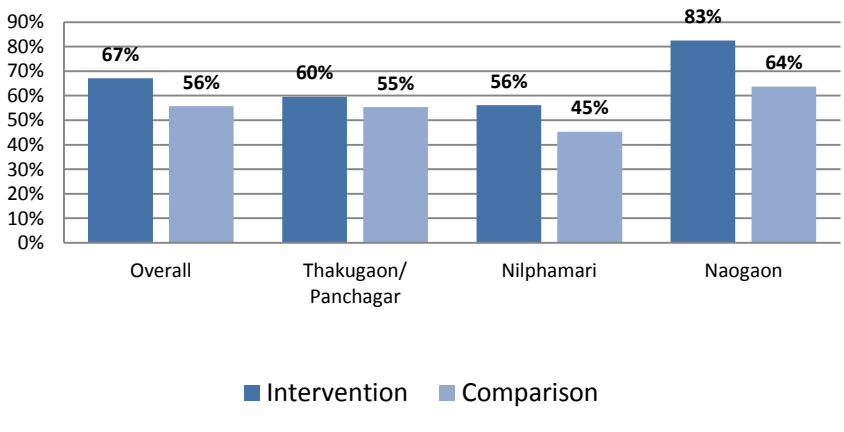
Surprisingly, half of the respondents in the Naogaon site reported themselves or their spouses as being change makers.

FIGURE 6.2.1:
Percentage of Surveyed Respondents Identifying Themselves or Spouses as Change Makers



The respondents were also asked whether they had been exposed to VAW messages through various media during the past two years. If the campaign had been implemented more intensely in Naogaon, we would expect differences to be reflected here as well. This is particularly with respect to interventions central to the campaign, e.g. personal interaction by change makers, campaigning, and various information, communication, and educational material distributed through the campaign, e.g. through pamphlets, leaflets, etc. The following four graphs present differences in reported exposure for these various media. As is apparent, significantly larger differences were observed for Naogaon in relation to all the media, but particularly for those central to the campaign. Differences also exist for Nilphamari, but these are less extreme.

FIGURE 6.2.2:
Percentage of Surveyed Respondents Reporting to Have Been Exposed to VAW Messages via Radio, TV, etc. in Last 2 Years



Exposure to various VAW messages was observed as being much greater in the Naogaon site.

FIGURE 6.2.3:
Percentage of Surveyed Respondents Reporting to Have Been Exposed to VAW Messages via Personal Interaction in Last 2 Years

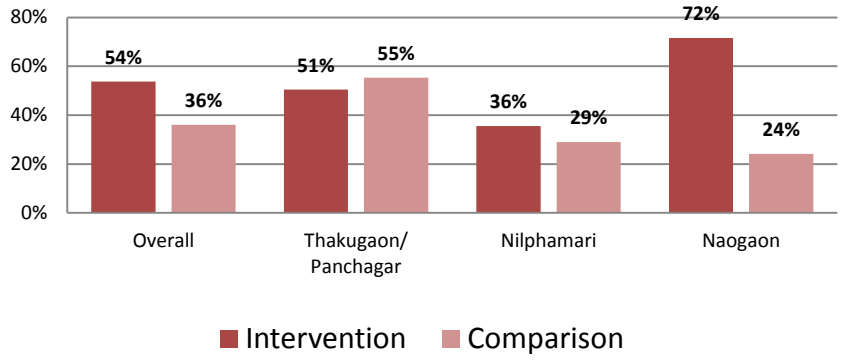


FIGURE 6.2.4:
Percentage of Surveyed Respondents Reporting to Have Been Exposed to VAW Messages via Campaign in Last 2 Years

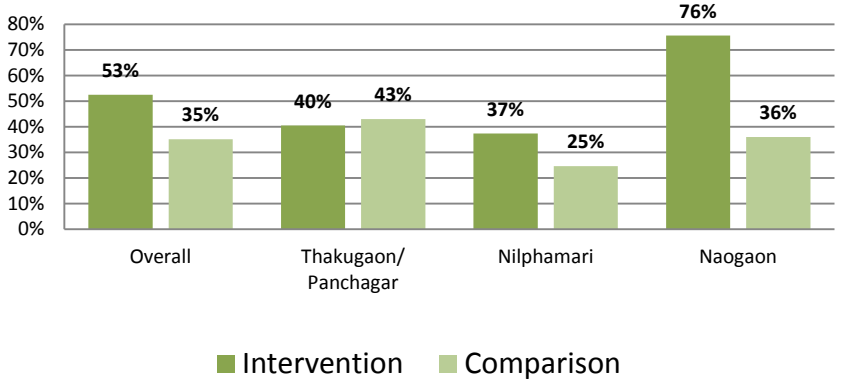
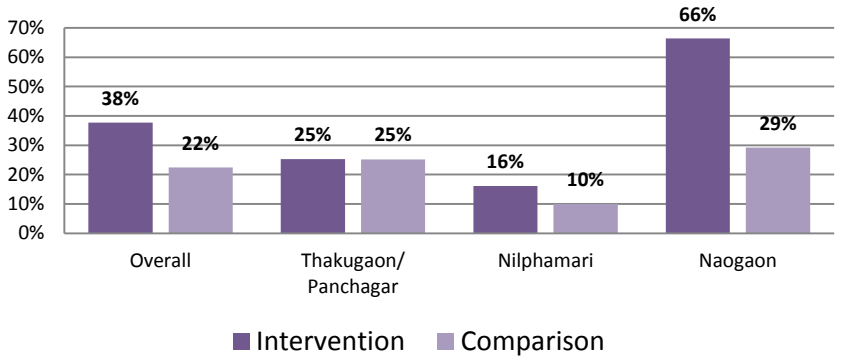


FIGURE 6.2.5:
Percentage of Surveyed Respondents Reporting to Have Been Exposed to VAW Messages via Printed Material in Last 2 Years



6.3 Differences Between the Intervention and Comparison Sites on the Outcome Measures

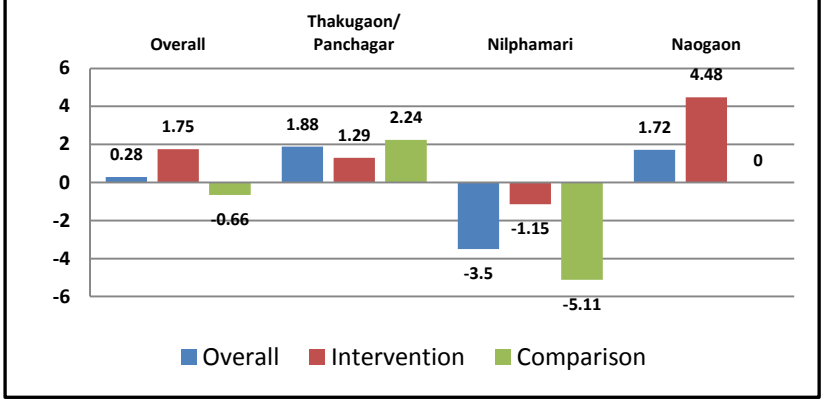
This subsection presents the results of analyses that compared the respondents from the intervention and comparison sites in relation to the outcome measures presented in Subsection 3.3.

6.3.1 Gender Attitudes

Overall, the respondents were not found to have a particularly negative or positive gender attitudes. However, differences between the districts were identified.

Figure 6.3.1.1 presents a graph of scores derived from the administration of the gender attitudes module as presented above. Recall that the maximum score a respondent could achieve was three for each question and the lowest score was -3. Given that there are 10 questions, the highest possible total score is 30 and the lowest score is -30. As is apparent from the graph, the overall average score is close to 0. This would imply that the average respondent, on the whole, has a neutral attitude. The overall raw score is furthermore better in the intervention site than the comparison site, and this also applies to Niphamari and Naogaon.

FIGURE 6.3.1.1:
Differences in Raw Gender Attitude Scores Between Intervention and Comparison Respondents



As mentioned in Subsection 3.3, factor analysis was used to better narrow in on the variation shared by the respondent's responses, and these factor scores formed the basis of the statistical analysis that was undertaken. The results of this analysis are presented in Table 6.3.1.1. Overall, respondents from the intervention group possess better attitudes, and the statistical significance of this difference holds for the two PSM and three MVR estimation procedures. However, the effect sizes generated by MVR are much larger and more statistically significant than those generated by PSM.

The picture is very different at district level. There is no positive statistically significant difference in the Thakurgaon/Panchagar site. The unadjusted difference of the Nilphamari site is statistically significant, but this statistical significance is inconsistent across the five estimation producers. All the effect estimates for Naogaon are highly statistically significant.

We can formally test whether the apparent differences between the intervention sites are statistically significant by carrying out an interaction test. Here, dummy interaction variables were generated by interacting each site with the intervention dummy variable. These dummy variables were then included the first MVR model that was used, i.e. the first MVR model presented in Table 6.3.1.1. Stata's *test* command was then used to see if the coefficients associated with the interacted terms are statistically different using the Wald test for interaction. The results are presented in Table 6.3.1.2 below, clearly indicating that the estimated effects of the campaign are statistically different across the intervention sites.

Both men and women residing in the Naogaon intervention villages were identified as have significantly more positive gender attitudes than their comparators.

TABLE 6.3.1.1:
Comparison of Intervention and Comparison Sites in Relation to Gender Attitudes (Principal Factor Score)

	Overall	Thakurgaon/ Panchagar	Nilphamari	Naogaon
Unadjusted:				
Sample mean	-0.00	0.00	-0.00	-0.00
Intervention mean:	0.13	-0.02	0.19	0.20
Comparison mean:	-0.08	0.01	-0.13	-0.13
Unadjusted difference :	0.208***	-0.0322	0.322***	0.331***
	(5.67)	(-0.53)	(4.68)	(5.41)
Observations:	**2247**	**750**	**644**	**853**
PSM (ATT)				
Post-matching difference:	0.128**	-0.129	0.0770	0.405***
(kernel)	(2.98)	(-1.92)	(1.01)	(5.67)
Observations:	**2094**	**746**	**515**	**833**
Post-matching difference:	0.135**	-0.0209	0.142	0.342***
(no replacement)	(3.02)	(-0.27)	(1.81)	(4.18)
Observations:	**2008**	**721**	**494**	**793**
Multivariable Regression:				
MVR coefficient (fe; robust):	0.237***	-0.0414	0.159	0.464***
	(5.47)	(-0.55)	(1.77)	(6.60)
Observations:	**2247**	**750**	**644**	**853**
MVR coefficient (fe; rreg):	0.217***	-0.108	0.184	0.448***
	(5.01)	(-1.52)	(1.87)	(6.02)
Observations:	**2247**	**750**	**644**	**853**
MVR coefficient (fe; robust):	0.221***	-0.103	0.140	0.453***
with control functions	(4.56)	(-1.26)	(1.32)	(6.11)
Observations:	**2246**	**744**	**623**	**852**

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
PSM estimates bootstrapped 1000 repetitions
Coefficients for covariates used not presented

TABLE 6.3.1.2:
Results of Survey Site Interaction Test for Gender Attitudes Measure Regressed on Intervention Dummy Variable

Original MVR Coefficient	MVR Coefficient with siteXintervention interaction variables	Wald Interaction Test (*F* statistic)
0.237***	0.462***	14.03***
(5.47)	(6.94)	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
coefficients for covariates not presented

Given that a fairly large effect estimate was identified for the Naogaon site, coupled with the non-experimental nature of the data, it is of interest to explore how much unobserved bias would be needed to "explain away" the effect. In other words, just how sensitive is the Naogaon effect estimate to the possibility of some unobserved and, by extension, uncontrolled for difference(s) between the respondents intervention and comparison villages?

Sensitivity analysis is an approach used for exploring this. It was implemented using Rosenbaum sensitivity analysis[4] with Stata's *rbounds* command. Here, unobserved bias is assumed to exist among both members of the intervention and comparison group at different log odds ratios. How large can the odds ratio be in order to render the effect estimate in question non-significant? Table 6.3.1.3 presents the results that were obtained from undertaking such analysis with the nearest neighbour one-to-one matching effect estimate. The

table reveals that the presence of unobserved bias would need to be present at a log odds ratio of 1.4 in favour of the intervention population in order for the effect estimate to be rendered statistically insignificant. Qualitatively, we can say that the effect estimate is moderately, but not strongly, robust to the possible existence of omitted variable bias.

TABLE 6.3.1.3:
Results of Rosenbaum Sensitivity Analysis Where Unobserved, Positive Bias is Assumed to Exist a Various Odds Ratios Among the Intervention Population in the Naogaon Site

Log odds ratio of hidden bias	*p*-value of effect estimate with bias	Estimated effect with bias	95% confidence level – two tailed CI+	CI−
1	.000039	.320287	.164773	.485092
1.1	.000514	.266763	.10909	.540596
1.2	.003716	.217879	.059926	.594788
1.3	.016835	.173882	.012779	.641107
1.4	.053127	.132735	-.028395	.684493
1.5	.126387	.091958	-.067016	.725412
1.6	.240809	.057814	-.104814	.767065
1.7	.385478	.022509	-.138188	.802698
1.8	.53888	-.00799	-.169888	.836268
1.9	.679175	-.036545	-.200805	.868476
2	.792355	-.064276	-.229638	.90148

It is of obvious interest to observe the extent the picture changes when the data are disaggregated by the sex of the respondent. Table 6.3.1.2 presents the results of the relevant analyses that were undertaken. The effect estimates generated by the various estimation procedures are more variable in this case. Perhaps the results of greatest interest are those associated with the Naogaon site. Almost all the effect estimates are statistically significant for both the female and male respondents, but those of the latter appear considerably larger.

In the Naogaon site, men's gender attitudes appear to have been more positively impacted by the campaign. However, an interaction test did not reveal this to be statistically significant.

TABLE 6.3.1.3
Comparison of Intervention and Comparison Sites in Relation to Gender Attitudes – Female/Male (Principal Factor Score)

	Overall		Thakurgaon/ Panchagar		Nilphamari		Naogaon	
	F	M	F	M	F	M	F	M
Unadjusted:								
Sample mean	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00
Intervention mean:	0.11	0.14	-0.02	-0.02	0.16	0.22	0.19	0.22
Comparison mean:	-0.07	-0.09	0.01	0.01	-0.11	-0.15	-0.12	-0.13
Unadjusted difference :	0.186***	0.231***	-0.0297	-0.0348	0.273**	0.374***	0.310***	0.352***
	(3.55)	(4.48)	(-0.31)	(-0.46)	(2.99)	(3.61)	(3.61)	(4.03)
Observations:	**1134**	**1113**	**380**	**370**	**331**	**313**	**423**	**430**
PSM (ATT)								
Post-matching difference:	0.165*	0.0923	-0.104	-0.154	0.260**	-0.129	0.338**	0.465***
(kernel)	(2.55)	(1.56)	(-1.02)	(-1.79)	(2.80)	(-1.04)	(2.91)	(4.79)
Observations:	**1068**	**1026**	**376**	**370**	**289**	**226**	**403**	**430**
Post-matching difference:	0.150*	0.161*	-0.0601	-0.0194	0.274**	-0.0535	0.283*	0.449***
(no replacement)	(2.41)	(2.56)	(-0.55)	(-0.19)	(2.87)	(-0.43)	(2.50)	(4.08)
Observations:	**1037**	**971**	**376**	**345**	**289**	**205**	**372**	**421**
Multivariable Regression:								
MVR coefficient (fe; robust):	0.156*	0.310***	-0.118	-0.0497	0.188	0.171	0.231*	0.543***
	(2.49)	(5.48)	(-1.04)	(-0.58)	(1.45)	(1.47)	(2.14)	(5.67)
Observations:	**1134**	**1113**	**380**	**370**	**331**	**313**	**423**	**430**
MVR coefficient (rreg):	0.122*	0.279***	-0.158	-0.0960	0.169	0.215	0.230*	0.519***
	(1.97)	(4.99)	(-1.36)	(-1.16)	(1.18)	(1.60)	(2.04)	(5.10)
Observations:	**1134**	**1113**	**380**	**370**	**331**	**312**	**423**	**430**
MVR coefficient (robust):	0.119	0.301***	-0.141	-0.169	-0.119	0.471***	0.195	0.564***
with control functions	(1.72)	(4.71)	(-1.10)	(-1.84)	(-0.84)	(3.60)	(1.81)	(5.63)
Observations:	**1134**	**1111**	**376**	**365**	**312**	**300**	**423**	**427**

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
PSM estimates bootstrapped 1000 repetitions
Coefficients for covariates used not presented

We can formally test the extent these differences are statistically significant by carrying out another interaction test. This time the sex of the respondent was interacted with the dummy intervention variable and included in the MVR model used specifically for the Naongaon site. The results are presented in Table 6.3.1.3. Surprisingly, the differences in the effect sizes estimated for the female and male respondents are not statically significant.

TABLE 6.3.1.3:
Results of Respondent Sex Interaction Test for Gender Attitudes Measure Regressed on Intervention Dummy Variable for Naogaon Site

Original MVR Coefficient	New MVR Coefficient with interaction variable	Coefficient for sexXintervention variable
0.464***	0.452***	0.025
(6.60)	(4.94)	(0.20)

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
coefficients for covariates not presented

Recall from Subsection 6.2, a good number of the respondents (approximately eight percent overall) were change makers. It is also of interest to assess if the impacts of the campaign are different for this particular subgroup. However, given that only very few of the respondents in the comparison sites reported themselves to be change makers and the mobilisation of change makers is a key intervention of the campaign, performing interaction tests is not useful. Another approach is to examine how the effect estimates of the campaign on gender attitudes changes when a dummy variable indicating whether the respondent is a change maker or not is included in the model. If there are no unique effects of the campaign on change makers, we would expect the overall coefficient for the intervention dummy variable to remain unchanged and the coefficient for the change maker variable to be statistically insignificant.

The estimated effects of the campaign on gender attitudes changes significantly with the inclusion of the dummy change maker variable in the analysis.

TABLE 6.3.1.4:
Results of MRV Analysis with Inclusion of Change Maker Dummy Variable

	Original Intervention Coefficient	Intervention site coefficient with CM dummy	Change maker coefficient	Post estimation Wald test (F-statistic)
Overall	0.237***	0.103*	0.551***	34.78***
	(5.47)	(2.24)	(6.64)	
Thakurgaon/ Panchagar	-0.0414	-0.044	0.055	0.19
	(-0.55)	(-0.58)	(0.29)	
Nilphamari	0.159	0.148	0.418	3.80*
	(1.77)	(1.63)	(1.93)	
Naogaon	0.453***	0.174	0.576	35.46***
	(6.11)	(2.12)	(5.71)	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
coefficients for covariates not presented

Table 6.3.1.4 presents the results of the relevant analysis. The original overall and Naogaon coefficients change significantly with the inclusion of the change maker dummy variables, and the coefficients for these particular variables are highly statistically significant. The extent the differences between the new

intervention and the change maker coefficients are different was formally tested. The results of the tests are in the last column. Statistically significant differences were identified for all sites, save for Thakurgaon/Panchagar.

The gender attitudes of the change makers, then, seem to have been more greatly impacted by the campaign in comparison with other women and men in the intervention sites. It is of interest, therefore, to explore whether there is any evidence that the campaign made any impact at all on the gender attitudes of non-change makers in the intervention sites. This is very relevant, given that the We Can Campaign theory of change assumes that the change makers will influence non-change makers in the communities. One way of exploring this is simply to analyse the data with the exclusion of all respondents who identified themselves as change makers from the analysis. Table 6.3.1.5 presents the results of MVR analysis that was undertaken to this end. While the overall and Naogaon specific effect sizes are substantially smaller, there are, nonetheless, statistically significant. There is, therefore, evidence that the impact of the campaign has extended beyond the change maker subgroup, particularly in Naogaon district.

The estimated impact of the campaign on gender attitudes still remains significant, particularly, in the Naogaon site, when change makers are excluded from the analysis .

TABLE 6.3.1.5:
Results of MVR Analysis of Comparison of Intervention and Comparison Sites in Relation to Gender Attitudes with Exclusion of Change Makers

	Overall	Thakurgaon/ Panchagar	Nilphamari	Naogaon
Multivariable Regression:				
MVR coefficient (fe; robust):	0.0997*	-0.0477	0.107	0.287***
	(2.16)	(-0.63)	(1.21)	(3.41)
Observations:	**2060**	**736**	**626**	**698**

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Coefficients for covariates used not presented

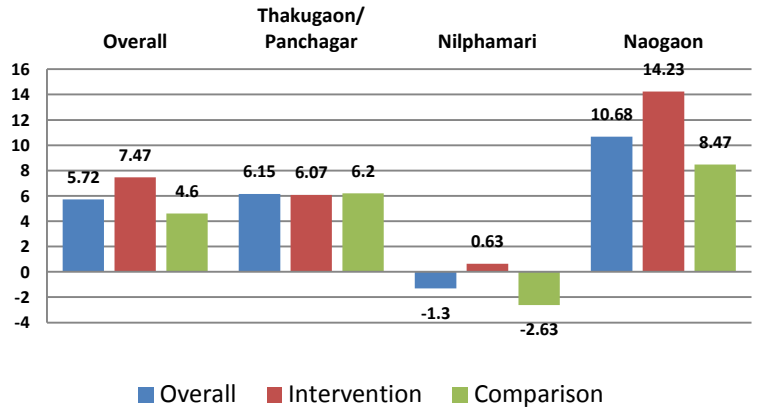
6.3.2 VAW Attitudes

Recall from Subsection 3.3 that data pertaining to two attitudinal measures of VAW were collected. Figure 6.3.2.1 presents a graph of raw scores derived from the administration of the Likert style measure. This particular scale is comprised of 12 items. As such, the maximum possible score a respondent could obtain was 36 and the lowest was -36. As indicated in the graph, most of the scores are positive, save for Nilphamari, indicating that most of the respondents were more likely to respond is a desirable way to the various statements. There is a small difference in the overall scores for the intervention and comparison groups, and there is considerable variability among the three research sites.

As was the case with general gender attitudes, the main statistical analyses that were undertaken were based on the factor score, rather than raw score, of this particular VAW attitudinal measure. Table 6.3.2.1 presents the results of the various analytical procedures that were implemented. The results are similar to those of the general gender attitudes measure: Overall, the results are positive, but this appears primarily due to the influence of the strong effect estimates derived for the Naogaon site. There is no evidence of impact in the Thakurgaon/Panchagar site, both before and after the implementation

of the statistical adjustment procedures. The unadjusted difference between the intervention and comparison sites for Nilpharmari is large, but this disappears following the implementation of PSM and MVR.

FIGURE 6.3.2.1:
Differences in Raw VAW Attitude Scores Between Intervention and Comparison Respondents



There is evidence that the campaign positively affected VAW attitudes in the Naogaon site but not the others.

TABLE 6.3.2.1:
Comparison of Intervention and Comparison Sites in Relation to Intra-marital Violence Attitudes (Principal Factor Score)

	Overall	Thakurgaon/ Panchagar	Nilphamari	Naogaon
Unadjusted:				
Sample mean	-0.00	-0.00	-0.00	-0.00
Intervention mean:	0.15	0.01	0.15	0.27
Comparison mean:	-0.09	-0.01	-0.10	-0.17
Unadjusted difference :	0.242***	0.0177	0.250***	0.431***
	(6.32)	(0.27)	(3.50)	(7.05)
Observations:	**2234**	**742**	**645**	**847**
PSM (ATT)				
Post-matching difference:	0.199***	0.0477	-0.00930	0.510***
(kernel)	(4.06)	(0.58)	(-0.11)	(6.11)
Observations:	**2064**	**733**	**508**	**823**
Post-matching difference:	0.181***	0.0338	0.0345	0.410***
(no replacement)	(3.96)	(0.42)	(0.42)	(5.19)
Observations:	**1980**	**709**	**488**	**783**
Multivariable Regression:				
MVR coefficient (fe; robust):	0.251***	0.0601	-0.0954	0.429***
	(5.84)	(0.73)	(-1.05)	(6.19)
Observations:	**2234**	**742**	**645**	**847**
MVR coefficient (rreg):	0.238***	0.0255	-0.0982	0.421***
	(5.24)	(0.30)	(-0.97)	(5.76)
Observations:	**2234**	**742**	**645**	**847**
MVR coefficient (robust):	0.230***	0.120	-0.0567	0.379***
with control functions	(4.99)	(1.30)	(-0.51)	(5.30)
Observations:	**2233**	**736**	**620**	**846**

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
PSM estimates bootstrapped 1000 repetitions
Coefficients for covariates used not presented

We can further formally test whether the differences in the effect sizes among the sites are indeed statistically significant by implementing a Wald test for

interaction. The results are presented in Table 6.3.2.2, clearly revealing differences in the estimated impacts of the programme on VAW attitudes in the sites.

TABLE 6.3.2.2:
Results of Survey Site Interaction Test for VAW Attitudes Measure Regressed on Intervention Dummy Variable

Original MVR Coefficient	MVR Coefficient with siteXintervention interaction variables	Wald Interaction Test (*F* statistic)
0.251***	0.469***	11.50***
(5.84)	(7.25)	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
coefficients for covariates not presented

Given the size of the effect estimate obtained for the Naogaon site on the VAW measure, sensitivity analysis was implemented again. The results are presented in Table 6.3.2.3. This time unobserved, positive bias would need to be present at an odds ratio of 1.5 in order to explain away the estimated effect. Again, this reveals that the effect estimate is moderately robust to the possible presence of omitted variable bias.

A fair degree of bias would be needed to explain away the estimated effect of the campaign on VAW attitudes in Naogaon district.

TABLE 6.3.2.3:
Results of Rosenbaum Sensitivity Analysis Where Unobserved, Positive Bias is Assumed to Exist a Various Odds Ratios among the Intervention Population in the Naogaon Site

Log odds ratio of hidden bias	*p*-value of effect estimate with bias	Estimated effect with bias	95% confidence level – two tailed CI+	CI−
1	.000028	.329281	.172361	.493324
1.1	.000385	.274662	.116253	.551924
1.2	.002891	.225868	.065993	.603545
1.3	.013565	.181082	.019916	.650064
1.4	.044254	.140026	-.021647	.693739
1.5	.108591	.100774	-.060331	.735411
1.6	.21289	.064672	-.097607	.776541
1.7	.349707	.03054	-.132399	.810602
1.8	.500167	-.000164	-.164019	.844692
1.9	.642832	-.028862	-.194733	.876967
2	.762122	-.057457	-.224937	.910724

It is again of interest to examine whether the effects differ when the data are disaggregated by the sex of the respondent. The results of the gender disaggregated analyses are presented in Table 6.3.2.4 below. Again, there appears to be a difference in the effect sizes for the Naogaon site between the male and female respondents. And this time the effect sizes appear to be larger for women than men. Another interaction test was undertaken to assess whether the apparent difference, at least for the first MVR output, is statistically significant. The results are presented in Table 6.3.2.5, and reveal that the difference is indeed statistically significant.

While there is evidence that both women and men's VAW attitudes were impacted by the campaign in the Naogaon site, the impact on women was greater.

TABLE 6.3.2.4:
Comparison of Intervention and Comparison Sites in Relation to Intra-marital Violence Attitudes – Female/Male (Principal Factor Score)

	Overall		Thakurgaon/ Panchagar		Nilphamari		Naogaon	
	F	M	F	M	F	M	F	M
Unadjusted:								
Sample mean	-0.00	-0.00	-0.00	-0.00	-0.00	0.00	-0.00	-0.00
Intervention mean:	0.17	0.12	0.01	0.02	0.17	0.13	0.33	0.20
Comparison mean:	-0.11	-0.07	-0.00	-0.01	-0.12	-0.09	-0.21	-0.12
Unadjusted difference :	0.288***	0.194***	0.00970	0.0260	0.283**	0.216*	0.544***	0.321***
	(5.53)	(3.46)	(0.11)	(0.27)	(3.03)	(1.98)	(6.46)	(3.62)
Observations:	**1124**	**1110**	**378**	**364**	**330**	**315**	**416**	**431**
PSM (ATT)								
Post-matching difference:	0.286***	0.113	-0.0609	0.161	0.221*	-0.266*	0.689***	0.353***
(kernel)	(4.12)	(1.61)	(-0.59)	(1.33)	(2.26)	(-2.02)	(4.91)	(3.42)
Observations:	**1055**	**1009**	**374**	**359**	**287**	**221**	**394**	**429**
Post-matching difference:	0.242***	0.128	-0.0398	0.107	0.218*	-0.222	0.567***	0.411***
(no replacement)	(4.03)	(1.88)	(-0.36)	(0.88)	(2.09)	(-1.83)	(4.84)	(3.91)
Observations:	**1024**	**956**	**374**	**335**	**287**	**201**	**363**	**420**
Multivariable Regression:								
MVR coefficient (fe; robust):	0.265***	0.195**	-0.0817	0.175	-0.0760	-0.136	0.538***	0.323**
	(4.48)	(3.19)	(-0.70)	(1.51)	(-0.55)	(-1.11)	(5.36)	(2.99)
Observations:	**1124**	**1110**	**378**	**364**	**330**	**315**	**416**	**431**
MVR coefficient (rreg):	0.227***	0.181**	-0.164	0.213	-0.0598	-0.167	0.464***	0.316**
	(3.76)	(2.79)	(-1.40)	(1.76)	(-0.39)	(-1.22)	(4.65)	(2.87)
Observations:	**1124**	**1110**	**378**	**364**	**330**	**315**	**416**	**431**
MVR coefficient (robust):	0.216***	0.215**	-0.0528	0.336**	-0.192	0.0590	0.468***	0.323**
with control functions	(3.39)	(3.29)	(-0.39)	(2.73)	(-1.08)	(0.44)	(4.62)	(2.87)
Observations:	**1124**	**1108**	**374**	**359**	**312**	**302**	**416**	**428**

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
PSM estimates bootstrapped 1000 repetitions
Coefficients for covariates used not presented

TABLE 6.3.2.5:
Results of Respondent Sex Interaction Test for VAW Attitudes Measure Regressed on Intervention Dummy Variable for Naogaon Site

Original MVR Coefficient	MVR Coefficient with sexXintervention interaction variables	Coefficient for sexXintervention
0.429***	0.281	0.306*
(6.19)	(2.94)	(2.48)

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
coefficients for covariates not presented

Have the impacts of the campaign on VAW attitudes been greater among the change makers than for the general population of the intervention sites? To answer this question, MVR analyses were again undertaken by including the dummy change maker variable. The results are presented in Table 6.3.2.6. It is clear that the inclusion of the change maker dummy variable does significantly change the estimated effects of the programme on VAW attitudes. This is the case overall and for two out of the three survey sites.

TABLE 6.3.2.6:
Results of MRV Analysis on Effects of Campaign with Inclusion of Change Maker Dummy Variable

	Original Intervention Coefficient	Intervention site coefficient with CM dummy	Change maker coefficient	Post estimation Wald test (F-statistic)
Overall	0.251***	0.096	0.634	67.81***
	(5.84)	(2.04)	(9.18)	
Thakurgaon/ Panchagar	0.0601	0.011	0.760	7.71***
	(0.73)	(0.14)	(3.88)	
Nilphamari	-0.0954	-0.101	0.188	1.02
	(-1.05)	(-1.10)	(0.96)	
Naogaon	0.429***	0.100	0.654	55.60***
	(6.19)	(1.13)	(7.14)	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Coefficients for covariates not presented.

It is of further interest to observe what happens when the change makers are excluded from the analysis, to assess whether the campaign may have changed VAW attitudes among non-change makers. The results of MVR analyses undertaken to this end are presented in Table 6.3.2.7. The results are very interesting and relevant: Now neither the overall nor the Naogaon specific effects are statistically significant, indicating that the campaign has done little to improve VAW attitudes among non-change makers.

When change makers are excluded from the analysis, the identified effect of the campaign on the first VAW attitudinal measure disappears.

TABLE 6.3.2.7:
Results of MVR Analysis of Comparison of Intervention and Comparison Sites in Relation to VAW Attitudes with Exclusion of Change Makers

	Overall	Thakurgaon/ Panchagar	Nilphamari	Naogaon
Multivariable Regression:				
MVR coefficient (fe; robust):	0.0918	0.0229	-0.0931	0.143
	(1.89)	(0.27)	(-1.00)	(1.53)
Observations:	**2048**	**728**	**627**	**693**

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Coefficients for covariates used not presented

Do we get similar results with the second VAW attitude measure? Recall that this measure involved asking whether a husband has good reason to hit his wife depending specific scenarios. An answer of "yes" to any of these scenarios indicates that the respondent condones inter-marital violence, at least in certain circumstances. The results of the analyses that were undertaken are presented in Table 6.3.2.8. Overall, 81 percent of the respondents appear to find it acceptable for a husband to hit his wife in specific situations. However, there is a desirable and statistically significant difference between the intervention and comparison groups. Overall and for the Naogan site, the statistical significance of these effect estimates are robust to the various PSM and MVR procedures that were undertaken. It is clear, again, that there are differences between the sites, and the figures presented in Table 6.3.2.9 confirm this.

There is an overall desirable effect for the second attitude measure, but – again – this due the influence of the Naogaon site .

TABLE 6.3.2.8:
Comparison of Intervention and Comparison Sites in Relation to Intra-marital Violence Attitudes – Measure 2 (Binary Outcome)

	Overall	Thakurgaon/ Panchagar	Nilphamari	Naogaon
Unadjusted:				
Sample mean	0.81	0.81	0.89	0.76
Intervention mean:	0.76	0.79	0.87	0.65
Comparison mean:	0.85	0.82	0.91	0.82
Unadjusted odds ratio:	-0.0883***	-0.0307	-0.0474	-0.1777***
(*exponentiated coefficients*)	(-5.22)	(-1.04)	(-1.92)	(-5.79)
Observations:	**2266**	**752**	**656**	**858**
PSM (ATT)				
Post-matching difference:	-0.0860***	-0.0284	-0.0499	-0.179***
(kernel)	(-4.40)	(-1.05)	(-1.94)	(-4.08)
Observations:	**2213**	**752**	**652**	**809**
Post-matching difference:	-0.0762***	-0.0246	-0.0570*	-0.170***
(no replacement)	(-4.05)	(-0.84)	(-2.05)	(-3.89)
Observations:	**2135**	**752**	**652**	**731**
Multivariable Regression:				
Adjusted odds ratio (fe; robust):	-0.1082***	-0.0277	-0.0111	-0.2685***
	(-5.75)	(-0.95)	(-0.27)	(-6.72)
Observations:	**2265**	**740**	**613**	**857**
Adjusted odds ratio:	-0.1075***	-0.0175	-0.0151	-0.2567***
with control functions	(-5.15)	(-0.59)	(-0.55)	(-6.21)
Observations:	**2265**	**740**	**589**	**857**

t statistics in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
PSM estimates bootstrapped 1000 repetitions
coefficients for covariates not presented

TABLE 6.3.2.9:
Results of Survey Site Interaction Test for Binary VAW Attitudes Measure Regressed on Intervention Dummy Variable

Original MVR Coefficient	MVR Coefficient with siteXintervention interaction variables	Wald Interaction Test (*F* statistic)
0.452***	0.266***	13.48**
(-5.75)	(-6.57)	

t statistics in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
coefficients for covariates not presented

It is of further interest to explore how robust the effect estimate obtained for the Naogaon site is to being explained away by possible unmeasured differences between the intervention and comparison respondents. Sensitivity analysis was therefore again carried out. The results are presented in Table 6.3.2.10. As indicated, the odds ratio of unobserved bias would need to be nearly twice as prevalent among the respondents in the intervention group in order to explain away the significance of the effect.

Are there differences in the effects of the campaign among women and men on this particular attitudinal measure? Table 6.3.2.10 presents the results of the gender disaggregated PSM and MVR analyses that were undertaken. The results again reveal quite a large difference between women and men. This is both the case overall and for the Naogaon site. Overall, for instance, the female respondents from the intervention sites are about half as likely as their female comparators to say that a husband sometimes has good reason to hit his wife, while men are a little over a quarter as likely. The differences appear even greater in the Naogaon site. However, interaction tests were carried out test whether these differences are indeed statistically significant, and this is, again,

surprisingly not the case (see Table 6.3.2.11).

The effect estimate obtained for the Naogaon site is considerably robust to being explained away by unmeasured and uncontrolled for bias.

TABLE 6.3.2.10:
Results of Rosenbaum Sensitivity Analysis Where Unobserved, Positive Bias is Assumed to Exist a Various Odds Ratios Among the Intervention Population in the Naogaon Site

Odds Ratio of Hidden Bias	Estimated effect with bias (odds ratio)	*p*-value of effect estimate with bias
1		.000019
1.1	3.75697	.000094
1.2	3.34535	.000353
1.3	2.9681	.001057
1.4	2.62317	.002654
1.5	2.30543	.005787
1.6	2.01084	.011253
1.7	1.73622	.019916
1.8	1.47897	.032598
1.9	1.23699	.049968
2	1.0085	.072456

The apparent differential gendered impact of the campaign in the Naogaon site was found to be statistically insignificant.

TABLE 6.3.2.10:
Comparison of Intervention and Comparison Sites in Relation to Intra-marital Violence Attitudes – Measure 2 (Binary Outcome)

	Overall		Thakurgaon/ Panchagar		Nilphamari		Naogaon	
	F	M	F	M	F	M	F	M
Unadjusted:								
Sample mean	0.79	0.83	0.66	0.96	0.90	0.89	0.83	0.68
Intervention mean:	0.74	0.78	0.64	0.94	0.85	0.88	0.74	0.55
Comparison mean:	0.83	0.87	0.67	0.97	0.93	0.90	0.89	0.76
Unadjusted odds ratio:	0.602***	0.530***	0.862	0.517	0.429*	0.879	0.364***	0.385***
	(-3.45)	(-3.94)	(-0.67)	(-1.25)	(-2.32)	(-0.36)	(-3.84)	(-4.49)
Observations:	**1139**	**1127**	**379**	**373**	**334**	**322**	**426**	**432**
PSM (ATT)								
Post-matching difference:	-0.0730*	-0.10***	-0.0406	-0.0160	-0.0825*	-0.0139	-0.101	-0.24***
(kernel)	(-2.32)	(-4.25)	(-0.80)	(-0.64)	(-2.32)	(-0.38)	(-1.36)	(-4.67)
Observations:	**1094**	**1119**	**379**	**373**	**334**	**318**	**381**	**428**
Post-matching difference:	-0.0614*	-0.079**	-0.0139	-0.0071	-0.094**	0.0160	-0.100	-0.22***
(no replacement)	(-2.13)	(-3.11)	(-0.27)	(-0.31)	(-2.78)	(0.40)	(-1.62)	(-3.76)
Observations:	**1035**	**1100**	**379**	**373**	**334**	**318**	**322**	**409**
Multivariable Regression:								
Adjusted odds ratio (fe; robust):	0.562**	0.271***	0.918	0.268	1.246	0.539	0.402*	0.117***
	(-2.96)	(-5.23)	(-0.25)	(-1.09)	(0.28)	(-0.75)	(-2.07)	(-5.31)
Observations:	**1131**	**1125**	**375**	**251**	**321**	**287**	**424**	**431**
Adjusted odds ratio:	0.634*	0.255***	0.927	0.503	0.881	0.152	0.906*	0.128***
with control functions	(-2.13)	(-4.98)	(-0.21)	(-0.49)	(-0.16)	(-1.45)	(-2.13)	(-5.01)
Observations:	**1131**	**1125**	**375**	**251**	**306**	**276**	**426**	**429**

t statistics in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
PSM estimates bootstrapped 1000 repetitions

TABLE 6.3.2.11:
Results of Respondent Sex Interaction Test for VAW Attitudes Measure Regressed on Intervention Dummy Variable for Naogaon Site

	Original MVR Odds Ratio	Overall MVR Odds Ratio with sexXintervention interaction variables	Odds ratio for sexXintervention interaction term
Overall	0.452***	0.428***	1.11
	(-5.75)	(-4.59)	(0.42)
Naogaon	0.204***	0.221	0.824
	(-6.72)	(-5.46)	(-0.49)

t statistics in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
coefficients for covariates not presented

Are the effects of the campaign on this second VAW attitudinal measure also restricted to the change makers? The dummy change maker variable was again included in the logistic MVR model, and the results are presented in Table 6.3.2.12. It is again clear that the campaign has had a considerably greater impact on the change makers in relation to this attitudinal measure.

What happens when the change makers are removed from the analyses? As is presented in Table 6.3.2.13, the overall odds ratio is now no longer statistically significant at with a 95 per cent level of confidence. However, that associated with the Naogaon site remains statistically significant, providing some evidence that the impacts of the campaign on attitudes has spilled over to non-change makers.

When change makers are excluded from the analysis, the identified second VAW attitudinal effect is reduced but remains statistically significant.

TABLE 6.3.2.12:
Results of MRV Analysis on Effects of Campaign on Binary VAW Attitude Measure with Inclusion of Change Maker Dummy Variable

	Original Intervention odds ratio	Intervention site odds ratio with CM dummy	Change maker odds ratio	Post estimation Wald test (F-statistic)
Overall	0.452***	0.747	0.195***	90.73***
	(-5.75)	(-1.92)	(-7.65)	0.0000
Thakurgaon/ Panchagar	0.751	0.839	0.070	10.63**
	(-0.95)	(-0.58)	(-3.11)	
Nilphamari	0.867	0.918	0.284	3.02
	(-0.27)	(-0.16)	(-1.73)	
Naogaon	0.204***	0.556*	0.160	77.00***
	(-6.72)	(-2.07)	(-6.13)	

z statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
coefficients for covariates not presented

TABLE 6.3.2.13:
Results of MVR Analysis of Comparison of Intervention and Comparison Sites in Relation to Binary VAW Attitude Measure with Exclusion of Change Makers

	Overall	Thakurgaon/ Panchagar	Nilphamari	Naogaon
Multivariable Regression:				
MVR coefficient (fe; robust):	-0.0343	-0.0103	-0.0053	-0.1113**
	(-1.84)	(-0.42)	(-0.02)	(-2.61)
Observations:	**2079**	**727**	**595**	**693**

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Coefficients for covariates used not presented

6.3.3 Women's Experience of Intra-marital Violence

1. Any Type of Physical Intra-marital Violence:

Given that one of the core aims of the We Can Campaign is to reduce intra-marital violence, this section of the report is of obvious interest. Table 6.3.3.1 presents the results of analysis that compared the intervention and comparison women in relation to self-reported intra-marital physical violence of any kind

over the last 12 months. As is evident, close to 60 percent of all interviewed married women reported being subjected to such violence. The overall unadjusted difference between the intervention and comparison groups is quite small. When disaggregated, the difference for Nilphmari is considerable. However, the statistical significance of this difference is not consistent following the PSM and MVR procedures that were undertaken. These results provide little in the way of evidence that the We Can Campaign model has reduced inter-marital violence in the study sites.

Following statistical adjustment, there are no consistent differences between the intervention and comparison sites in relation the physical VAW experience.

TABLE 6.3.3.1:
Comparison of Intervention and Comparison Sites in Relation to Experience of Marital Physical Violence in Last 12 Months (Binary Outcome)

	Overall	Thakurgaon/ Panchagar	Nilphamari	Naogaon
Unadjusted:				
Sample mean	0.59	0.62	0.72	0.47
Intervention mean:	0.57	0.64	0.62	0.45
Comparison mean:	0.61	0.60	0.80	0.48
Unadjusted probit coefficient:	-0.0467	0.0371	-0.171***	-0.0315
	(-1.57)	(0.72)	(-3.45)	(-0.63)
Observations:	**1139**	**380**	**334**	**425**
PSM (ATT)				
Post-matching difference:	-0.0136	0.0444	-0.106	0.0107
(kernel)	(-0.41)	(0.86)	(-1.82)	(0.20)
Observations:	**1132**	**380**	**329**	**423**
Post-matching difference:	-0.0161	0.0138	-0.118*	-0.0185
(no replacement)	(-0.50)	(0.24)	(-2.11)	(-0.34)
Observations:	**1126**	**380**	**323**	**423**
Multivariable Regression:				
Adjusted probit coefficient:	-0.0555	0.0184	-0.0780	-0.005
(fe; robust)	(-1.52)	(0.26)	(-0.98)	(-0.03)
Observations:	**1139**	**376**	**330**	**423**
Adjusted odds ratio:	-0.0556	0.069	-0.176*	-0.0462
with control functions	(-1.38)	(0.85)	(-1.96)	(-0.60)
Observations:	**1139**	**376**	**319**	**423**

z statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
PSM estimates bootstrapped 1000 repetitions

It is important to note that the results presented in Table 6.3.3.1 pertain to all the women that were interviewed. Moreover, recalled baseline data pertaining to intra-marital violence experience were not included in the analysis, as this would have resulted in the dropping out of those respondents who had married since the baseline period.

Does the picture change with the inclusion of the recalled baseline data? Table 6.3.3.2 displays the results of the relevant analyses. The two overall MVR coefficients are now statistically significant, as is the case for the Naogaon site. However, the effect estimates generated from the Niphamari site still remain inconsistent.

Given that both follow-up and baseline data exist for many of the respondents, double difference effect estimates were also computed. These are presented in Table 6.3.3.3. Recall that the married female respondents were asked whether they had been subjected to any physical violence, both in the last 12 months and back in 2004. If a respondent reported that they had been subjected to such violence back in 2004 but had not in the last 12 months, they were coded

with 1 and 0 otherwise. In other words, a variable was created for women who reported experiencing a cessation of violence since the baseline.

With the inclusion of baseline data, there is some evidence that the campaign reduced VAW in the Naogaon.

TABLE 6.3.3.2:
Comparison of Intervention and Comparison Sites in Relation to Experience of Marital Physical Violence in Last 12 Months (with Baseline Physical Violence Control Variable)

	Overall	Thakurgaon/ Panchagar	Nilphamari	Naogaon
Unadjusted:				
Sample mean	0.57	0.59	0.73	0.44
Intervention mean:	0.52	0.61	0.58	0.41
Comparison mean:	0.61	0.58	0.82	0.46
Unadjusted probit coefficient:	-0.0863*	0.0308	-0.2449***	-0.059
	(-2.36)	(0.48)	(-4.01)	(-1.00)
Observations:	**767**	**239**	**231**	**297**
PSM (ATT)				
Post-matching difference:	-0.0708	0.0408	-0.124	-0.127
(kernel)	(-1.74)	(0.63)	(-1.91)	(-1.86)
Observations:	**753**	**238**	**225**	**290**
Post-matching difference:	-0.0688	0.0208	-0.155*	-0.0917
(no replacement)	(-1.65)	(0.30)	(-2.34)	(-1.30)
Observations:	**738**	**238**	**210**	**290**
Multivariable Regression:				
Probit coefficient (fe; robust):	-0.1324*	0.007	-0.0206	-.2584*
	(-2.53)	(0.07)	(-0.22)	(-2.26)
Observations:	**767**	**238**	**229**	**295**
Probit coefficient:	-0.1307*	-0.0326	-0.0575	-0.239*
with control functions	(-2.28)	(-0.29)	(-0.55)	(-1.98)
Observations:	**767**	**232**	**215**	**295**

z statistics in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
PSM estimates bootstrapped 1000 repetitions

The double difference effect estimate is significant for Naogaon across all estimation procedures.

TABLE 6.3.3.3:
Comparison of Intervention and Comparison Sites in Relation to Reported Cessation of Subjection to Any Physical Violence Since Baseline (binary double difference)

	Overall	Thakurgaon/ Panchagar	Nilphamari	Naogaon
Unadjusted:				
Sample mean	0.16	0.23	0.10	0.16
Intervention mean:	0.20	0.22	0.12	0.25
Comparison mean:	0.14	0.23	0.09	0.10
Unadjusted dif. :	0.066*	-0.016	0.037	.0151***
	(2.39)	(-0.29)	(0.91)	(3.40)
Observations:	**767**	**239**	**231**	**297**
PSM (ATT)				
Post-matching difference:	0.0487	-0.0707	0.0383	0.155***
(kernel)	(1.57)	(-1.04)	(0.91)	(3.37)
Observations:	**744**	**237**	**231**	**276**
Post-matching difference:	0.0707*	-0.0330	0.0778	0.138**
(no replacement)	(2.42)	(-0.55)	(1.85)	(2.88)
Observations:	**740**	**233**	**231**	**276**
Multivariable Regression:				
Adjusted dif. (fe; robust):	0.075*	0.001	-0.052	0.096*
	(2.53)	(0.07)	(-0.90)	(2.41)
Observations:	**767**	**239**	**222**	**297**
Adjusted dif:	0.075*	0.021	-0.067	0.094*
with control functions	(2.25)	(0.33)	(-1.08)	(1.98)
Observations:	**750**	**214**	**210**	**269**

z statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
PSM estimates bootstrapped 1000 repetitions

Overall, 20 percent of the respondents reported no longer being subjected to intra-marital violence in the intervention sites in comparison with 14 percent comparison sites. This overall difference remains statistically significant across three of the four estimation procedures. Moreover, all the adjusted differences remain statistically significant for the Naogaon site.

Is the effect estimate identified for Naogaon statistically different from the other sites? Again, the relevant interaction test was carried out. The results are presented in Table 6.3.3.4, revealing statistically significant differences in the effects of the campaign in the various sites. The data, therefore, provide some evidence that the campaign reduced intra-marital violence for the average woman residing in the Naogaon intervention site who had been married at the campaign's baseline period.

TABLE 6.3.3.4:
Results of Survey Site Interaction Test for Binary Cessation of Violence Measure Regressed on Intervention Dummy Variable

Original MVR Coefficient	MVR Coefficient with siteXintervention interaction variables	Wald Interaction Test (*F* statistic)
0.075*	0.1889***	8.56*
(2.53)	(3.88)	

z statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
coefficients for covariates not presented

Sensitivity analysis was again carried out to assess how much omitted variable bias would be needed to explain away the effect identified in the Naogaon site. Given the binary nature of the outcome variable, Mantel-Haenszel bounds analysis was undertaken, rather than Rosenbaum bounds analysis. The effect estimates generated by this procedure are displayed in the form of an odds ratio. The results are presented in Table 6.3.3.5 below. As indicated in the table, the odds of such bias would need to be present at an odds ratio of 1.4 to render the effect estimate statistically insignificant, indicating that the identified effect is moderately robust to hidden bias.

It would take a moderate amount of unobserved bias to explain away the effect identified in Naogaon.

TABLE 6.3.3.5:
Results of Rosenbaum Sensitivity Analysis Where Unobserved, Positive Bias is Assumed to Exist a Various Odds Ratios Among the Intervention Population in the Naogaon Site

Odds Ratio of Hidden Bias	Estimated effect with bias (odds ratio)	*p*-value of effect estimate with bias
1	2.55209	0.005354
1.1	2.27967	0.011314
1.2	2.02846	0.021257
1.3	1.79926	0.035988
1.4	1.58851	0.056086
1.5	1.39341	0.081747
1.6	1.21177	0.1128
1.7	1.04181	0.14875
1.8	0.882072	0.188869
1.9	0.731364	0.232278
2	0.588681	0.278038

Even with the removal of change makers and spouses who are change makers from the analysis, the effect estimate obtained for Naogaon still holds.

How does the apparent impact of the campaign on the cessation of violence on some of the women residing in intervention sites change when change makers or women who are married to change makers are excluded from the analysis? Table 6.3.3.6 presents the results of MVR analyses that were undertaken. When either female change makers or women married to change makers are taken out of the analysis, the overall and Naogaon specific effects of the campaign still remain statistically significant. There is evidence, therefore, that the campaign has reduced intra-marital violence among the general population in the Naogaon site.

TABLE 6.3.3.6:
Results of MVR Analysis of Comparison of Intervention and Comparison Sites in Relation to Reported Cessation of Violence with Exclusion of Change Makers

	Overall	Thakurgaon/ Panchagar	Nilphamari	Naogaon
MVR without Change Makers:				
MVR coefficient (fe; robust):	0.0768*	0.0027	-0.0233	0.1742**
	(2.25)	(0.23)	(-0.43)	(2.87)
Observations:	**684**	**232**	**219**	**233**
MVR without Husband Change Makers:				
MVR coefficient (fe; robust):	0.0776*	0.0025	-0.0491	0.1184*
	(2.49)	(0.21)	(-0.85)	(2.51)
Observations:	**684**	**235**	**225**	**270**

z statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Coefficients for covariates used not presented

In conclusion, there is evidence to suggest that the campaign has brought about a reduction in general physical violence but only in Naogaon survey site. This effect, however, was not detected when the interviewed women were compared without controlling for baseline differences in violence experience. This is because the women in the Naogaon intervention villages reported being subjected to more intra-marital violence in the baseline period than their comparators. It was only after controlling for this difference that the apparent effect of the campaign was derived.

2. Serious Physical Intra-marital Violence:

Does the picture change when we narrow in on more serious forms of violence? Recall from Subsection 3.3 that this would be where the married female respondent reported being punched with a closed fist, seriously beaten, and so forth. Table 6.3.3.6 presents the results of the various analyses that were performed in relation to this particular outcome variable for all the interviewed women. No statistically significant difference was identified between the intervention and comparison sites, particularly after implementation of the statistical adjustment procedures.

How does the situation change when we narrow in on those women who were married at the time of the campaign's baseline survey? Table 6.3.3.7 present the results of the analyses that were undertaken in relation to differences in the percentages of women who reported the cessation of serious intra-marital violence since the baseline period. Again, none of the adjusted effect estimates are statistically significant.

TABLE 6.3.3.7:
Comparison of Intervention and Comparison Sites in Relation to Experience of Serious Marital Physical Violence in Last 12 Months (Binary Outcome)

	Overall	Thakurgaon/ Panchagar	Nilphamari	Naogaon
Unadjusted:				
Sample mean	0.43	0.51	0.56	0.25
Intervention mean:	0.39	0.50	0.49	0.21
Comparison mean:	0.45	0.51	0.61	0.28
Unadjusted difference :	-0.0579	-0.0072	-0.1144*	-0.0686
	(-1.93)	(-0.14)	(-2.07)	(-1.59)
Observations:	**1139**	**380**	**334**	**425**
PSM (ATT)				
Post-matching difference:	-0.0220	-0.0234	-0.0152	-0.0265
(kernel)	(-0.74)	(-0.42)	(-0.26)	(-0.62)
Observations:	**1136**	**380**	**333**	**423**
Post-matching difference:	-0.0322	-0.00690	-0.0547	-0.0741
(no replacement)	(-1.05)	(-0.12)	(-0.88)	(-1.62)
Observations:	**1127**	**380**	**324**	**423**
Multivariable Regression:				
Probit Coefficient (fe; robust):	-0.0604	-0.075	-0.136	-0.0268
	(-1.63)	(-1.04)	(-1.30)	(-0.47)
Observations:	**1139**	**378**	**333**	**423**
Probit coefficient	-0.0273	0.0363	-0.1699	-0.0233
with control functions:	(-0.66)	(0.44)	(-1.42)	(-0.41)
Observations:	**1139**	**378**	**319**	**423**

z statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
PSM estimates bootstrapped 1000 repetitions

No statistically significant effect of the campaign was identified in relation to serious violence, even after controlling for baseline violence experience.

TABLE 6.3.3.8:
Comparison of Intervention and Comparison Sites in Relation to Reported Cessation in Experience of Serious Physical Intra-marital Violence Since Baseline (binary double difference)

	Overall	Thakurgaon/ Panchagar	Nilphamari	Naogaon
Unadjusted:				
Sample mean	0.16	0.23	0.10	0.14
Intervention mean:	0.16	0.20	0.08	0.20
Comparison mean:	0.15	0.26	0.11	0.10
Unadjusted coefficient :	0.01091	-0.0647	-0.0286	0.0988*
	(0.41)	(-1.16)	(-0.72)	(2.37)
Observations:	**767**	**239**	**231**	**297**
PSM (ATT)				
Post-matching difference:	0.0106	-0.0623	-0.0120	0.0879
(kernel)	(0.33)	(-1.08)	(-0.23)	(1.58)
Observations:	**680**	**237**	**167**	**276**
Post-matching difference:	0	-0.0842	0	0.108
(no replacement)	(0.00)	(-1.48)	(0.00)	(1.84)
Observations:	**625**	**237**	**154**	**234**
Multivariable Regression:				
Adjusted odds ratio (fe; robust):	0.0306	-0.0052	0.0098	0.0301
	(1.07)	(-0.26)	(0.25)	(0.97)
Observations:	**767**	**236**	**227**	**297**
Adjusted odds ratio:	0.0209	-0.0041	-0.0374	0.0238
with control functions	(0.66)	(-0.16)	(-2.19)	(0.74)
Observations:	**767**	**232**	**215**	**297**

z statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
PSM estimates bootstrapped 1000 repetitions

3. Mental Intra-marital Violence:

The picture would not be complete if we failed to examine whether there is any evidence that the We Can Campaign has reduced women's subjection to mental violence, as opposed to physical violence. Recall that this relates to whether the husband of the respondent had insulted or humiliated her in the last 12 months. Table 6.3.3.8 presents the results for all women who were interviewed. Nearly 40 percent reported being subjected to such violence and there is little difference between the intervention and comparison groups, particularly after the implementation of PSM and MVR.

How does the situation change when the reported drop in being subjected to mental violence since the baseline period is examined? Table 6.3.3.9 presents the results of analysis that compared percentages of women who reported that their husbands had not insulted or humiliated them in the last 12 months, but had done so during the baseline period. Here, several of the campaign effect estimates are statistically significant, both overall and for the Naogaon site, but this is not consistent across the estimation procedures.

There is some evidence that the campaign may have brought about a reduction in mental violence in the Naogaon site.

TABLE 6.3.3.8:
Comparison of Intervention and Comparison Sites in Relation to Experience of Intra-marital Mental Violence in Last 12 Months (Binary Outcome)

	Overall	Thakurgaon/ Panchagar	Nilphamari	Naogaon
Unadjusted:				
Sample mean	0.37	0.39	0.40	0.33
Intervention mean:	0.38	0.38	0.36	0.38
Comparison mean:	0.37	0.40	0.43	0.30
Unadjusted difference :	0.005	-0.021	-0.0714	0.0864
	(0.17)	(-0.40)	(-1.31)	(1.84)
Observations:	**1140**	**380**	**334**	**426**
PSM (ATT)				
Post-matching difference:	-0.0213	-0.0626	-0.0399	0.0302
(kernel)	(-0.65)	(-1.14)	(-0.64)	(0.57)
Observations:	**1096**	**378**	**329**	**389**
Post-matching difference:	0.00693	-0.00690	-0.0242	0.0183
(no replacement)	(0.22)	(-0.13)	(-0.41)	(0.36)
Observations:	**1087**	**378**	**320**	**389**
Multivariable Regression:				
Probit Coefficient (fe; robust):	0.0138	-0.0268	-0.120	0.0177
	(0.39)	(-0.38)	(-1.37)	(0.29)
Observations:	**1140**	**376**	**326**	**426**
Probit coefficient	-0.0102	-0.055	-0.1130	0.0213
with control functions:	(-0.26)	(-0.71)	(-1.04)	(0.33)
Observations:	**1140**	**376**	**312**	**426**

z statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
PSM estimates bootstrapped 1000 repetitions

TABLE 6.3.3.9:
Comparison of Intervention & Comparison Sites in Relation to Reported Cessation of Experience of Mental Intra-marital Violence Since Baseline(binary double difference)

	Overall	Thakurgaon/ Panchagar	Nilphamari	Naogaon
Unadjusted:				
Sample mean	0.13	0.15	0.10	0.14
Intervention mean:	0.15	0.15	0.09	0.20
Comparison mean:	0.12	0.15	0.10	0.10
Unadjusted coefficient :	0.0333	-0.0003	-0.0104	0.0933*
	(1.33)	(-0.01)	(-0.26)	(2.22)
Observations:	**767**	**239**	**231**	**297**
PSM (ATT)				
Post-matching difference:	0.0403	-0.0106	0.0341	0.0878*
(kernel)	(1.48)	(-0.19)	(0.89)	(1.99)
Observations:	**703**	**239**	**167**	**297**
Post-matching difference:	0.0249	0	0.0260	0.0431
(no replacement)	(1.00)	(0.00)	(0.60)	(0.97)
Observations:	**681**	**230**	**154**	**297**
Multivariable Regression:				
Probit coefficient (fe; robust):	0.066	0.0316	0.0046	0.0725
	(2.38)	(0.66)	(0.35)	(1.91)
Observations:	**767**	**239**	**231**	**293**
Probit coefficient	0.0687*	0.0623	-0.0075	0.07125
with control functions:	(2.36)	(0.95)	(-1.45)	(1.84)
Observations:	**767**	**232**	**215**	**297**

z statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
PSM estimates bootstrapped 1000 repetitions

6.3.4 Other Findings of Interest

Even with a more restricted comparison of villages with change makers, there is still no evidence of the campaign's impact in the Thakugaon/ Panchagar and Nilphamari sites.

- *Narrowing in on villages with change makers in Thakugaon/Panchagar and Nilphamari sites*

As presented in Subsection 6.2, implementation of the We Campaign in the villages of the Thakugaon/Panchagar and Nilphamari sites was not significantly intense. In particular, change makers were present in less than half of the villages and not much in the way of campaigning appears to have taken place in these villages. Given this variation in the implementation of the campaign in these sites, it is of interest to see if there is any evidence that it made an impact in particular villages where more complete implementation took place. Unfortunately, given that campaign activities were only reported to have taken place in a relatively small number of the villages of these sites, the analysis had to be restricted to comparing people residing in villages with change makers in the Thakugaon/Panchagar and Nilpharmari sites, with those residing in the comparison villages from these same sites.

Table 6.3.4.1 presents the results of these more restricted analyses. As indicated in the table, people residing in villages with change makers in these sites are not more likely to have better VAW attitudes than those residing in the comparison villages. In addition, no statistically significant difference was found in relation to women's experience of violence.

TABLE 6.3.4.1:
Results of MVR Analysis of Respondents Residing in Villages of the Thakugaon/Panchagar and Nilphamari Sites with Change Makers and Those Residing in the Comparison Villages in Relation to VAW Measures

	Overall	Thakurgaon/ Panchagar	Nilphamari
1. Violence attitude (Likert)			
MVR coefficient (fe; robust):	0.0771	0.00620	-0.0257
	(1.07)	(0.06)	(-0.24)
Observations:	**1090**	**610**	**480**
2. Violence attitude (binary)			
Probit coefficient (fe; robust):	-0.0262	-0.0528	0.007
	(-1.14)	(-1.56)	(0.27)
Observations:	**1084**	**607**	**459**
3. Reported Physical Violence			
Probit coefficient (fe; robust):	-0.0684	-0.0749	-0.0204
	(-1.18)	(-0.86)	(-0.28)
Observations:	**554**	**307**	**245**

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Coefficients for covariates used not presented

- *Does more exposure to VAW messages through various media have a greater impact on attitudes?*

As reported in Section 6.2, data were collected from the respondents on their exposure to VAW messages through the general media, one-to-one personal interaction, community-based campaigning, and print materials. Is there any evidence that more intense exposure to messages through these various media has a greater effect on attitudes? Is more better? To answer this question dose-response analysis was carried out for increasing levels of exposure to receipt of VAW messages through different types of media. The focus was on the intervention villages only to see if there is indeed a positive correlation between more intense exposure to such messages and VAW attitudes.

What are the most effective ways to deliver VAW messages to bring about attitudinal change?

To undertake the analysis the respondents were divided into four different exposure groups – no, low, medium, and high exposure – for each of the four media, namely one-to-one personal interaction, mass media, campaigning, and printed material. Two tests were then carried out to assess whether attitudes actually improve with increasing levels of reported exposure to various media – the linear contrast test and the departure from linear trend test. The former test assesses whether the attitude measure actually does increase in a statistically significant way with each increasing level of exposure. The departure from linear trend test, on the other hand, assesses if there is evidence that the relationship between level of exposure and the measure do not increase in a linear way. If the *F* statistic associated with the first test is statistically significant and that of the second test is not, then a clear linear relationship between exposure and the outcome measure is demonstrated.[5]

The results of the application of these two tests are presented in Table 6.3.4.2. All four types of media passed the linear contrast test, and all but the one-to-one interaction media passed the departure from linear trend test. While the effect of one-to-one interaction exposure may plateau at a certain level of intensity, the results of the analysis suggest that more intense exposure to all

the four media is correlated with better VAW attitudes.

Increased levels of exposure to all four media were correlated improved VAW attitudes.

TABLE 6.3.4.2:
Gradient MVR Coefficients and Results of Linear Contrast and Departure from Linear Trend Tests for VAW Attitudes Measure on 5 Levels of Various VAW Message Media – Intervention Site Only

Media	Level of exposure	MVR Coefficients	Linear Contrast Test (*F*-statistic)	Departure from Linear Trend Test (*F*-statistic)
VAW messages via one-to-one personal interaction	2	.2583637 (0.000)	55.89 (0.000)	3.73 (0.0244)
	3	.6169287 (0.000)		
	4	.5923759 (0.000)		
VAW messages via mass media	2	.1342471 (0.098)	50.33 (0.0000)	1.65 (0.1922)
	3	.3154652 (0.000)		
	4	.6500717 (0.000)		
VAW messages via campaigning	2	.2039101 (0.007)	46.92 (0.0000)	0.37 (0.6886)
	3	.3931522 (0.000)		
	4	.6989616 (0.000)		
VAW messages via print material	2	.2796547 (0.002)	37.02 (0.0000)	1.70 (0.1826)
	3	.2922557 (0.008)		
	4	.7004301 (0.000)		

p-values in parentheses

- *Further comparisons of the attitudes of change makers and non-change makers*

Disaggregating the data between male and female change makers also revealed some interesting findings. The relevant statistics are presented in Figure 6.3.4.1 and Figure 6.3.4.2 below. In the first figure, a comparison between change makers and non-makers in terms of condoning intra-marital VAW is presented. Not surprisingly, there are significant differences between the change makers and non-change makers. However, nearly half of all change makers still condone intra-marital VAW in specific circumstances. Even more surprising is that this statistic jumps to 60 per cent when female change makers are examined separately.

FIGURE 6.3.4.1:
Comparison of Change Makers with Non-change Makers with Respect to Condoning Intra-marital VAW (at Least in Certain Circumstances)

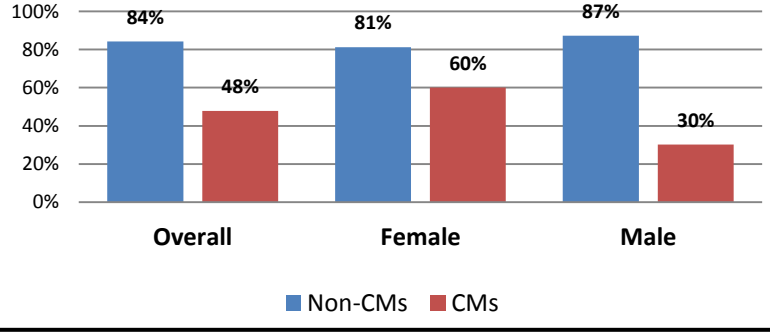
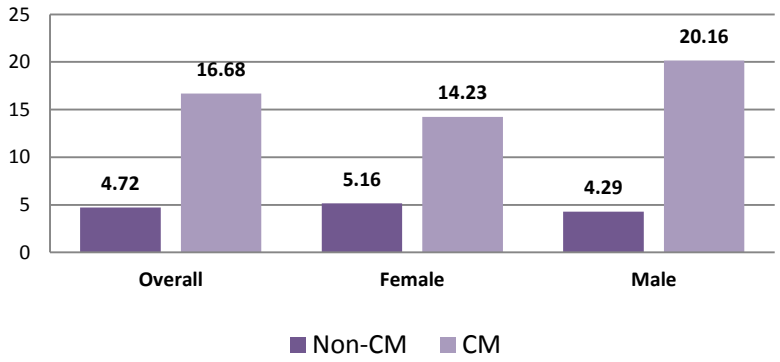


Figure 6.3.4.2 presents similar statistics but this time with the Likert scale attitude measure that was used. The picture is similar: There is a statistically significant difference between female and male change makers, with the latter having more positive attitudes than the former.

FIGURE 6.3.4.2:
Comparison of Change Makers with Non-change Makers with Respect to Raw VAW Attitude Score (Likert Scale)



7.0 Conclusions and Learning Considerations

There is fairly reliable evidence that the campaign brought about positive effects, but these are concentrated in the Naogaon site.

7.1 Conclusions

Overall, statistically significant and desirable differences were identified between the intervention and comparison populations in relation to general gender attitudes and, more specifically, attitudes pertaining to intra-marital violence. However, this is primarily due to the influence of the data that were collected from the intervention villages of the Naogaon district site. The differences found in this site are highly statistically significant and fairly robust to being explained away by the possible existence of unmeasured, confounding differences between the intervention and comparison populations that were not controlled for (i.e. unobserved heterogeneity).

Unfortunately, there is no evidence that the campaign has significantly impacted attitudes in the Thakugaon/Panchagar and Nilphamari sites. This was even the case when the analysis is only focused on those intervention villages that reportedly had change makers. This does not necessarily mean that the campaign had no impact on the change makers or those in the immediate sphere of influence in these districts. Such micro-level impacts may have taken place. However, the campaign had no wider detectable impact on the "average" man or woman in these communities.

It is also important to note that much of the campaign's impact on attitudes has been on the change makers themselves. In particular, when the change makers are removed from the analysis, the statistical significance of one of the measures is rendered insignificant, and the statistical significance of the others decreases considerably. Nevertheless, attitudes pertaining to the general gender attitude measure and the binary VAW measure still remained statistically significant in the more restricted analysis. There is, therefore, evidence that the campaign has positively affected attitudes in the

intervention villages of the Naogaon site beyond the change makers.

Naogaon is also the only site where there is some evidence that the campaign has brought about a reduction in women's experience of physical violence. The most reliable measure that was used was women's reported cessation of intra-marital violence since the baseline period. The various statistical adjustment procedures estimated that nine to 16 percent more women in the Naogaon intervention villages reported no longer experiencing such violence in comparison with women residing in this site's comparison villages.

Here and interestingly, the apparent impact is not greater among the change makers or even among women married to change makers compared to other women. The statistical significance of the difference holds even when change makers are removed from the analysis. There is, therefore, evidence that the campaign has brought about a reduction in intra-marital physical violence among the general population of married women in Naogaon site's intervention villages. This, of course, assumes that the recalled baseline data reported by the respondents in both the intervention and comparison sites are reliable.

The most plausible and likely explanation why there is evidence that the campaign brought about positive changes in the Naogaon site's intervention villages and not in those of the Thakugaon/Panchagar and Nilphamari sites is related to the campaign's implementation. There is strong evidence that the We Can Campaign was implemented more intensely and completely in the Naogaon site.

In conclusion, this cross-sectional impact assessment has found that the We Can Campaign model can bring about positive effects relevant for reducing violence experienced among women but only when implemented with a high level of intensity.

7.2 Programme Learning Considerations

Based on the findings of this effectiveness review, there are a number of points We Can Campaign stakeholders, both in Bangladesh and beyond, can consider to increase the effectiveness of the campaign.

A key finding is that implementation matters and it matters a lot! Ways to encourage intense implementation should be prioritised over geographic scale.

- *Identify key reasons for differences in campaign implementation in the Naogaon site vis-à-vis the Thakugaon/Panchagar and Nilphamari sites*

A key lesson learned is that implementation of the We Can Campaign matters. The data clearly revealed that implementation was significantly less intense in intervention villages surveyed in Thakugaon, Panchagar, and Niphamari districts, and this is the most plausible explanation why no differences in the outcome measures between the intervention and comparison populations of these districts were identified. Significant effects of the campaign were indentified for the Naogaon intervention villages, and the implementation of the campaign was much more intense in these villages.

Why did such differences in campaign implementation take place in the Naogaon site on the one hand and the Thakugaon/Panchagar and Niphamari

sites on the other? We came to understand that the implementation of the campaign took place later in the Naogaon site, i.e. from 2008 onwards. Could one of the explanations be that there is more local enthusiasm for the campaign because it is a more recent phenomenon? If this is the case, it has relevance for the sustainability of the campaign's impact.

Another possible explanation could be the Naogaon site benefiting from the lessons learned through the early experiences of implementing the campaign in the Thakugaon/Panchagar and Niphamari sites. Is a fundamentally different approach to implementing the campaign being pursued by *Polli Sree* in Naogaon? Are there any opportunities for replicating what is being done differently in Naogaon in the districts of Thakugaon, Panchagar, Niphamari, and perhaps even Dijnapur?

- *Explore ways to ensure that the We Can Campaign is carried out with significant intensity at the local level in the future, possibly sacrificing geographic scale*

Again, the findings of the effectiveness review revealed clearly that implementation matters. We also came to understand that significant efforts were made by *Polli Sree*, as well as the other We Can Alliance partners in Bangladesh, to mobilise increasing numbers of change makers to reach the campaign's targets. Could the pressure to expand the campaign to new areas have inadvertently resulted in lack of follow up with, as well as support to, previously mobilised geographic areas?

Given that the campaign seeks to changes perceptions, attitudes, and behaviour among not only the change makers but other members of the communities in which they reside, the findings of the review indicate that implementation intensity is critical. In this case, more intense – rather than simply more – is better.

The campaign may not work as per its hypothesised theory of change. Knowing its actual, as opposed to hypothesised, mechanism(s) would be useful.

- *Consider carrying out complementary qualitative research to interrogate, and possibly, challenge the We Can Campaign's theory of change*

The hypothesised way the campaign is to bring about one of its ultimate intended outcomes – reduction in intra-marital violence – was presented in Section 2.0. This is the campaign's "theory of change". While there is evidence to suggest that the campaign, where implemented with significant intensity, can reduce intra-martial violence, the findings of the review do raise some questions concerning the campaign's hypothesised mechanism.

Recall that the mechanism essentially boils down to one key hypothesis: Changing people's attitudes about VAW will change practice and, by extension, women's experience of VAW. As such, women come to realise that VAW is wrong and no longer tolerate it. Men also realise that hitting their wives, etc. is not a good thing, and therefore stop doing so.

While there is evidence that the campaign improved VAW attitudes among non-change makers in the intervention villages of the Naogaon sites, this was only in relation to one of the two VAW attitudinal measures. This indicates

that the impact of the campaign on the attitudes among non-change makers was not considerably significant. Had it been otherwise, we would have observed significant effects for both measures. Furthermore, for the one measure with a positive effect, the estimate would have been larger and more statistically significant.

Despite this, a greater reduction in intra-marital VAW was reported as having taken place among women residing in the Naogaon site's intervention villages, who are either not change makers themselves or married to change makers. It is quite possible that another mechanism triggered by the We Can Campaign brought about this impact. Imagine, for example, being a man who beats his wife who resides in a village where half of all married people are either change makers or married to change makers. Perhaps he may choose to change his behaviour not because he comes to understand that beating his wife is wrong but out of fear of being publically humiliated or stigmatised for being known as a wife beater.

So an alternative mechanism *may* have brought about a reduction in intra-marital VAW in the Naogaon intervention villages:

1. Critical mass of change makers mobilised in villages who regularly preach that VAW is wrong and husbands should stop hitting their wives.
2. Men fear being publically humiliated.
3. Men stop hitting their wives.

This is not to say at all that this is how the reduction in intra-marital VAW came about. It is only to suggest that another mechanism may have been at work. And, if so, it would be relevant for We Can Campaign stakeholders to know if this is the case. Better understanding of the actual, as opposed to hypothesised, mechanism(s), would allow modifications in the campaign's design to be made, so it better narrows in on triggering the real mechanism(s) at work. This would enable it to leverage more change.

Investing in in-depth, qualitative research to probe further into how the reduction in intra-marital VAW in the Naogaon site was actually brought about by the campaign would therefore be useful and help to strengthen the campaign's effectiveness.

- *Review either the design and/or implementation of the We Can "conscientisation" tools/processes among change makers.*

Why do so many female change makers condone intra-marital VAW?

A very interesting finding of the review is that approximately 60 percent of female change makers, as opposed to 30 percent of male change makers, reported that they find it justifiable for a man to hit his wife in particular circumstances. This is also true for the Naogaon site. If the change makers, as per the We Can theory of change, really do go through such a deep rooted attitudinal transformational process, how could this be possible?

This is also something that could be followed up with further qualitative research. Are, for example, many of the women signing up to be change makers not because they desire to end VAW or have been "conscientised", but because it is a popular thing for women in their villages to do? Or: Are

many women being issued with change maker identity cards who have not properly gone through "conscientisation" process? Or: Is there something wrong with the We Can change makers' tool kit itself and/or how it is being applied? All these questions can form the basis of further discussion and, possibly, follow-up research.

- *Investigate possibilities further testing the effectiveness of the We Can Campaign model*

While the findings of the effectiveness review provide some reasonably reliable evidence that the We Can Campaign model can – when implemented with a significant degree of intensity – change attitudes and women's experience of intra-marital VAW, further testing of its effectiveness and underlying mechanisms is recommended. The campaign was found to work in only one particular context – selected rural areas of Naogaon district. Are there particular contextual issues in these specific areas that enable it to work, or does it also work in other areas of Bangladesh and beyond? Would/does it work differently in urban areas?

Further impact assessment and research work would be useful to further test and strengthen the campaign's effectiveness.

Moreover, how intense does the implementation of the campaign need to be in order for it to be effective? If half the population of a village or municipality need to sign up as change makers in order for the campaign to bring about its effects, is this something that is feasible to replicate?

Finally, it must be fully acknowledged the findings of this effectiveness review are not as reliable as those that would have been obtained through a well managed randomised control trial, or even a more rigorous quasi-experimental design that included actual, as opposed to recalled, baseline data. Its results, however, do demonstrate that the We Can Campaign may very well have significant potential to reduce women's subjection to intra-marital violence, thereby making the pursuit of more involving impact evaluation designs worth investing in.

ANNEX 1: Scoring Key Using in OGB Supported Policy Influencing Effectiveness Reviews

Targeted Outcome	**Extent observed (high, medium, low, none)**	**Extent of project/campaign contribution (high, medium, low, none)**	**Specific contribution score*** **/5**	**Other evidenced explanations and extent of their contribution (high, medium, low)**
1.				•
2.				•
3.				•
4.				•
Unforeseen Outcome				
1.	n/a			•
2.	n/a			•

***Scoring Key – Specific Contribution of Project/Campaign**

Score	**Outcome Consideration**	**Contribution Consideration**
5 points	High level of outcome change realised	High project/campaign contribution
4 points	Medium level of outcome change realised	High project/campaign contribution
	High level of outcome change realised	Medium project/campaign contribution
3 points	Medium level of outcome change realised	Medium project/campaign contribution
	Low level of outcome change realised	High project/campaign contribution
2 points	High-medium outcome change realised	Low project/campaign contribution
	Low level of outcome change realised	Medium project/campaign contribution
1 point	Medium-low outcome change realised	Low project/campaign contribution
0 points	High-none outcome change realised	No project/campaign contribution
	Any negative unforeseen outcome change	High to low project/campaign contribution

REFERENCES

1. White H. A Contribution to Current Debates in Impact Evaluation. Evaluation. 2010 April 1, 2010;16(2):153-64.
2. Naved RT, Persson LA. Factors associated with spousal physical violence against women in Bangladesh. Studies in Family Planning. 2005;36(4):289-300.
3. Schuler SR, Islam F. Women's acceptance of intimate partner violence within marriage in rural Bangladesh. Studies in Family Planning. 2008;39(1):49-58.
4. Rosenbaum PR. Sensitivity Analysis in Observational Studies. In: Everitt BS, Howell DC, editors. Encyclopaedia of Statistics in Behavioral Science. Chichester: John Wiley & Sons; 2005.
5. Vittinghoff E, Shiboski S, Glidden D, McCulloch C. Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models: Springer; 2005.