



## Aligning Indicators and Ambitions: How to Improve Indicators Used in Programs to Reduce Violence Against Women and Girls

Todd Foglesong

### Introduction

Addressing violence against women and girls is among DFID's highest priorities, yet defining and measuring progress on this issue is especially difficult—the appropriate measures themselves elusive and the process of defining progress nettlesome. The temptation to elide these difficulties and borrow indicators used in programs vaguely reported to work elsewhere is considerable.

This paper is designed to help advisors avoid the use of quick and familiar indicators that too often prove empty. The paper provides practical suggestions for how DFID country offices might instead build and use locally generated indicators in Security & Justice programmes that are appropriate for measuring the results of activities designed to both reduce violence against women (VAWG) and improve the response to it.<sup>1</sup> Specifically, the paper:

- provides a framework that might guide the development of indicators in different contexts;
- suggests ways to strengthen indicators currently used in DFID programmes on VAWG;
- proposes solutions to key challenges in the collection and evaluation of relevant data.

This paper does not prescribe particular indicators. There are no turnkey indicators in this field. Instead, this paper recommends that the process of building indicators begin with an ambition defined both by

<sup>1</sup> The paper does not directly address indicators for social development programs, although the processes for indicator development described here may be helpful in that field, too. The paper recognizes that many DFID programs on violence against women are not stand-alone programs but rather small components of larger country development initiatives. Solutions to the design of indicators that harness synergies between programs are beyond the scope of this paper.

With funding from the United Kingdom's Department for International Development, (DFID), the Program in Criminal Justice Policy and Management (PCJ) at the Harvard Kennedy School has been supporting state officials and civil society organizations in Jamaica, Sierra Leone, and Nigeria to develop and use their own indicators to spark, reinforce, and communicate progress toward strategic goals in justice and safety. In 2010, PCJ began collaborating with officials in Papua New Guinea (PNG), extending existing efforts in the law and justice sector funded by the Australian Government Aid Program (AusAID).

The aim of the project is to equip government and civil society organizations with the skills and experience to design their own indicators, routinely assess those indicators, and use them to drive meaningful reform in the justice sector. Building this capacity is a long-term undertaking, for the desire for indicators and the skill in their construction must permeate the organizational culture in governmental and non-governmental bodies. It is also a fluid process: indicators serve ambitions, policies, governments, and staffs that inevitably change over time.

The prototype indicators developed in this project are different from the indicators in international systems created in the Global North for use in the Global South. They start by finding successes, however modest, and strengthen norms and standards that emerge in the course of reviewing local practices. They also perform different kinds of development work: They support domestic ambitions for justice and safety, reinforce management operations in government, and align the work of individual agencies with sector-wide goals. At the same time, these and other examples of country-led indicator development complement the growing number of globally conceived indicator projects by grounding the measurement culture of international development in local customs, and by articulating domestic sources of legitimacy for the standards implicit in the norms in global indicator projects.



**UKaid**  
from the Department for International Development



**Australian Government**  
AusAID

national governments and development programmes in a specific country and then articulate one or more theories of change that make explicit the processes by which DFID-supported activity could lead to the desired outcomes. It also recommends that indicators of progress be fashioned from locally available information which will test those theories of change. DFID advisors should assemble indicators from multiple sources of information according to the kinds of violence against women and girls they wish to reduce.

This paper begins by describing an unavoidable series of complications for this field and suggests a way through each of them. It then describes a set of choices that must be made in light of the gap between the scale of the ambition to end violence against women and girls and the means typically available to country programs for achieving that goal. Next it outlines a framework for indicator development, illustrating a process by which advisors might build indicators in support of a theory of change, using scenarios from current programming.

Finally, it reviews some of the existing indicators in several DFID country programs, proposing solutions to the challenges of assembling good measures from spotty data and imperfect sources of information.

### **Guidance Summary**

This paper suggests a number of solutions to practical challenges in the development of indicators for programs to reduce violence against women and improve the response to such violence. There is no established solution to these challenges. Indeed, this paper recommends experimentation in order to develop appropriate solutions since both the precise nature of the challenge and the character of possible solutions vary by country and development context. We list here some suggested approaches to six practical challenges that were described to us by governance advisors in several countries and are recognized by experts in the field.

<b>Practical Challenge</b>	<b>Suggested Approach(es)</b>	<b>Page</b>
Ambiguity of Prevalence Rates	Complement reported changes in prevalence rates with a second indicator to help interpret the meaning of the first	3
Parallel Ambitions for DFID and Country Partners	Develop measures that respond separately to the interests of national governments and DFID	5
Weak Evidentiary Base for Programming	Pair Measures of Output and Impact with Indicators of Outcomes	5
Potential Conflicts with Other Justice Goals	Warning indicators for perverse effects	6
Problems of Scale: Modest Means, Big Goals	Measure Proximate Changes; Develop Impact Multipliers; Focus on Intrinsic Values	9
Collecting, Managing, and Evaluating Data	Refine survey tools	16
	Supplement with administrative data	17
	Encourage local development of norms	18

## Table of Contents

1. Complications in Indicator Development for Programming on VAWG.....	3
2. Confronting Problems of Scale in Efforts to Reduce Violence Against Women and Girls.....	6
3. A Framework for Indicator Development in DFID Programming on VAWG.....	9
4. Development Scenarios, Theories of Change, and Justice Indicators.....	11
5. Indicators Currently Used in DFID Programming.....	13
6. Collecting, Managing, and Evaluating Data on VAWG.....	16
7. Conclusion.....	19

### 1. Complications in Indicator Development for Programming on VAWG

The design of good indicators on this subject is complicated by several familiar problems, among them (i) the ambiguity of prevalence measures, (ii) tensions between DFID's own ambitions to reduce VAWG and the approach of partners in national governments, and (iii) the weak evidentiary foundation for effective programming.

(i) Most people assume that good work on violence against women and girls will reduce its prevalence, but *effective programming can raise recorded prevalence rates* for at least two distinct reasons. First, good work on VAWG expands people's recognition of such violence, leading them to understand more behavior as violence. Especially with definitions of violence moving beyond physical violence to include economic and psychological violence, effective advocacy has the effect of expanding awareness and the likelihood of reporting in women's responses to surveys as well as interactions with state authorities.<sup>2</sup> Second, improved responses to VAWG can encourage more reporting, and this change could increase the prevalence rates recorded in surveys and also the number of crimes involving female victims that are registered by the government even if the amount of underlying violence is steady or declining.

<sup>2</sup> The second UN Special Rapporteur on Violence Against Women noticed this "paradox" in a speech to the Human Rights Council. "As States do more - change laws, increase protection, provide resources for NGO support services - it appears that levels of violence increase." See Yakin Erturk, A/HRC/7/6 January 29, 2008, p. 32, available at: <http://www.ohchr.org/EN/Issues/Women/SRWomen/Pages/AnnualReports.aspx>

The ambiguity of prevalence measures may be obvious, but it still confounds efforts to define and detect progress. After all, if an increase in levels of reported prevalence is a positive result, how should you interpret decreases in prevalence?

#### PRACTICE TIP

#### Coping with the Ambiguity of Prevalence Rates

Because an increase in the numbers of reports of violence can be interpreted either as a sign of growing confidence in the value of reporting or an indicator of more violence, you should always complement reported changes in prevalence rates with a second indicator to help interpret the meaning of the first. Here are three strategies you might consider:

- Ask respondents in the surveys that generate prevalence measures whether they think the incidence of violence against women is increasing, or just the awareness of such violence.
- Ask victims who request help from government or non-governmental organizations whether people they know have experienced more violence recently.
- Examine separately trends in grave and less grave violence, for if changes in grave violence are greater than those for less grave violence, the changes are likely to reflect prevalence. If changes are greater in less grave violence, those changes are more likely a sign of changes in willingness to report.

To assist in the interpretation of measures of prevalence, DFID advisors might develop a bundle of indicators, using multiple sources of information. Consider the following suggestions.

- *Measure shifts in all forms of violence*, looking for signs of changes in social norms. A survey in Punjab, India in 2010 discovered that 98 percent of women had experienced an “affront to their dignity by word, sound, or gesture intending to insult their modesty,” and yet only 4 incidents were reported to government agencies in 2010. An increase in the number of incidents of these types of violence reported to officials might indicate a reduction in tolerance for violence instead of an increase in their prevalence.
- *Locate proxy data for violence against women and girls*. Child sex ratios can be used as proxy indicators for female feticide and male child preference. They can also help manage the ambiguity of survey-based measures of prevalence. For example, the state of Bihar, India, recorded a decrease in the amount of violence against women in the years 2007-2010 and only a few incidents of female feticide. The state registered a reduction in the proportion of female children in the population in the same period. These kinds of public health and demographic data can help you interpret changes in police reports of violence against women and girls, distinguishing pockets of progress from persistent problems.
- *Compare modes and recipients of reporting*, distinguishing organizations that enjoy greater and lesser public trust. If, for example, reports of violence to the police remain constant while reports to other mediums of redress increase, it may be that rates of reporting rather than the number of incidents is on the rise. If, on the other hand, there is a decrease in the number of incidents reported to informal bodies such as women’s centers at schools and hospitals, or local political or religious bodies, or trusted specialized units for women and girls within

the police, it may reflect a genuine decrease in the amount of violence.<sup>3</sup>

*(ii) Tensions between DFID’s ambitions and those of national governments take countless forms*, even when there is a shared desire to increase protection of women and girls. Where DFID advisors might focus on women’s empowerment, a national government may prefer to focus on reinforcing family cohesion rather than disrupting patriarchal hierarchies. Where DFID might prefer a broad definition of violence against women, including psychological and economic violence, a national government might restrict the definition and measures to physical or sexual violence. These tensions may not be apparent in the choice of desired outcomes, but they will surface when theories of change are made explicit.

To manage this tension, you should develop at least one indicator that genuinely corresponds to the government’s stated ambition and a separate set of indicators that gauge the wider results DFID hopes will be achieved. By using the same sources of data to generate both measures you can reinforce a common interest in shared systems of information without forcing a false agreement on objectives.

*(iii) Country offices are forced to design programmes and indicators on a relatively weak evidentiary base*. Most of the available evidence about effective programming on VAWG comes from the Global North, where economic, cultural, and social conditions are very different from those where DFID is operating.<sup>4</sup> Moreover, even the most rigorously designed experiments yield mixed results: batterer rehabilitation programs, second responder systems, and mandatory arrest policies generate more violence in some families and communities and

---

<sup>3</sup> For more information about these examples and other insights about locally generated indicators of reductions in violence against women and children, see Rainuka Dagar, “Gender Based Violence: A Guide for Capacity Building of Gender Responsive Police Service Delivery,” Institute for Development and Communication in Chandigarh, India, [http://idcindia.org/index.php?option=com\\_content&view=article&id=100&Itemid=193](http://idcindia.org/index.php?option=com_content&view=article&id=100&Itemid=193)

<sup>4</sup> A briefing commissioned by DFID last year from the non-governmental organization, Saferworld, described these problems. See “Addressing Violence Against Women in Security and Justice Programmes,” March 11, 2010.

less in others.<sup>5</sup> Indicators that test locally generated hypotheses about how to reduce VAWG can help construct a genuinely global body of knowledge and evidence.

**PRACTICE TIP**

**Measuring Parallel Ambitions**

Indicators that effectively communicate in London the value of British investments abroad and indicators that serve the needs of governments in a developing country are rarely the same. DFID advisors always seek agreement as closely as possible in the overall program that is negotiated between governments, so it is all the more important that the tensions smoothed over in program design are respected through multiple indicators. As one DFID advisor put it, “one of the key tasks of indicators is to get the compromise right between what DFID wants, what the government can do, and what is best for the people.” The way you can use indicators to help manage the compromise throughout the life of a program is by assigning separate indicators to the objectives specific to the government and to DFID.

Consider using parallel measures. For example, at a shelter for women, you might:

- Use a prevalence indicator that encompasses the widest UN definition of violence.
- Help the police build an indicator for a category of violence important to the government.

To contribute to this knowledge, programs should develop *output, impact, and outcome measures*, especially where projects are testing new forms of prevention and assistance for women and girls. Managers of these interventions may be focused primarily on outputs they control (such as the quality of services in a shelter) and less interested in

questioning assumptions about their effectiveness. Special assistance might be required to develop impact and outcome measures that suit DFID’s purposes.

These are not the only complications in developing indicators for programming in violence against women and girls. As we describe in the sections below, the development of indicators in this field also must confront problems of scale, deficits in knowledge and limitations in the sources of data, and the thorny processes of defining problems together with government partners.

**PRACTICE TIP**

**Developing Indicators in the Absence of Strong Evidence**

Suppose you are supporting a shelter for victims of family violence where none existed before. How would you know whether the shelter reduces violence? At least three kinds of indicators could help you assess whether the shelter adds safety in the short term and also whether these same women experience more or less victimization after they leave:

- An *output indicator* might measure “successful protection” by counting the number of violent experiences of all women staying at the shelter, distinguishing between the types of problems for which they sought help and controlling for the duration of their stay.
- An *impact measure* to test the independent effect of the intervention might compare the experiences of two equally matched groups of women, one that spent time at the shelter and one that did not.
- An *outcome measure* to assess whether the shelter’s help changes lives might require some follow-up with a sample of these women, such as all victims that left the shelter last month, and inquire about the frequency and severity of recent experiences of violence.

<sup>5</sup> The World Health Organization recently concluded in its exhaustive summary of evidence in this field that: “the field of intimate partner and sexual violence prevention must therefore be considered to be at its earliest stages in terms of having an established evidence base.” See *Preventing Intimate Partner Violence and Sexual Violence Against Women: Taking Action and Generating Evidence*, World Health Organization, 2010, p. 37. For a recent academic discussion of the limits of this evidence, see “Experimental Studies of Violence Against Women,” a special issue of the *Journal of Experimental Criminology*, vol. 6, no. 4 (December 2010).

## **2. Confronting Problems of Scale in Efforts to Reduce Violence Against Women and Girls**

A central challenge in all development work is managing the gap between the scale of the ambition and the means available to achieve it. The gap between ends and means seems particularly acute in this field, where the global goal is to “end” or “eliminate” VAWG and yet national level investments are modest and individual interventions make minor dents in the problem as a whole. For example, even the most effective school-based education and prevention programs reach only a minority of communities in which VAWG may be a problem, and they sometimes capture the attention of individuals that are least likely to contribute to VAWG. To take another example, prosecutors will only have sufficient evidence to obtain convictions in a fraction of cases where police identify a perpetrator, which themselves constitute a small portion of all cases originally reported by victims. Even if police and prosecution become more effective, and none of the offenders that are held accountable by the courts reoffend, some quantum of VAWG will be unrelieved by the justice system.

To manage the problems of scale, DFID advisors might adjust both sides of the equation, stretching the capacities of systems of justice and security in modest ways, on the one hand, and overcoming shortcomings in the way goals in this field are defined and measured on the other.

### *Strengthening Systems of Justice and Safety*

Justice systems make modest contributions to the control of violence in any society, and they may be weak planks in the platforms on which governments construct a response to VAWG. Justice systems are designed to be retail operations, and even their best efforts in individual cases – dignified attention to victims, prolonged support for families, intensive work with offenders – are unlikely to generate wholesale reductions in the amount of crime and violence. The administration of justice in individual cases can contribute to better community safety, as some efforts to reduce repeat victimization of sexual assault, burglary, and robbery have shown.<sup>6</sup> But

---

<sup>6</sup> See, for example, Amanda Robinson, “Reducing repeat victimization among high-risk victims of domestic violence: the benefits of a coordinated community response in Cardiff, Wales,” *Violence Against Women*, v. 12, no. 8, August 2006. More recently, see “Supporting High-Risk Victims of Domestic Violence: A Review of Multi-Agency Risk Assessment Conferences (MARACs),” *Home Office Research Report*, 55, July

strategic reductions in violence are not customarily the concern of front-line justice workers. The imposition of such goals may also interfere with or seem to contradict other commitments, such as providing more personalized attention to individual victims and avoiding the multiple forms of “secondary victimization.” Leaders of justice institutions in the Global North themselves are only beginning to develop ways to align measures of quality and effectiveness.<sup>7</sup>

Burdening justice sector operations in developing countries with grand strategic goals and new international commitments could have corrosive effects. For example, asking young non-governmental organizations or inexperienced state agencies to deliver macro-level results could distract them from the management of basic operations, taxing limited capacities. Legal officials and justice service providers in newly democratic countries might also resent and resist the subordination of their operations to larger policy goals, especially if they come from interests, leaders, and policies that appear to be transient, or if they seem incompatible with aspirations for the “independence” of police, prosecutors, and judges. Finally, few managers of front-line operations possess the kind of information necessary to design and evaluate interventions that deliver strategic results, such as the reduction of repeat, serial, “series,” or “multiple” victimization.

Furthermore, not all of the goals DFID and other development organizations are pursuing in the justice sector are naturally compatible with metrics used in campaigns to end VAWG. For example, a program to improve criminal defendants’ access to effective legal representation or a program to reduce unlawful methods of police investigation may be strained by projects that seek to increase successful prosecution, especially if that goal is measured by the

---

2011. Available at: <http://www.homeoffice.gov.uk/publications/science-research-statistics/research-statistics/crime-research/horr55/horr55-report?view=Binary>

<sup>7</sup> Asked in November 2010 by the House of Commons Justice Committee about how to measure the contribution of the new Code of Practice for Victims to broader policy goals, the UK Commissioner for Victims and Witnesses replied: “I think not only is it not measurable but some of it just doesn’t add up.” See “Uncorrected transcript of oral evidence taken before the Justice Committee of the House of Commons, Tuesday 9 November 2010, available at: <http://www.publications.parliament.uk/pa/cm201011/cmselect/cmjust/uc583-i/uc58301.htm>

rate of conviction with or without trial.<sup>8</sup> Victims' ability to withdraw complaints after initially requesting prosecution might be impaired by projects to reduce the "rate of attrition" in prosecution, and some victims may be exposed to greater harm upon the prosecution and conviction of defendants. For these reasons, indicators of greater accountability for offenders must be balanced by measures for other important goals. Advisors might start small, with the development of measures for objectives such as greater respect, care, and dignity in the treatment of victims, or improvements in the quality of evidence gathered by investigators.

Another way to conceptualize the justice system's contribution to reducing violence against women and girls is to promote the visibility of justice operations, especially those that disseminate the message the violence is a crime not to be tolerated. In this way the justice system can both reiterate and broadcast the state's commitment to women's safety and security, and at the same time help to build trust in women and girls that the state will respond to violence.

### *Defining and Measuring Violence Against Women*

Standard measures of the prevalence of VAWG, including those recommended by the United Nations might not be suitable for DFID programs. Few country programs, for example, are likely to have substantial effects on the proportion of the female population between the age of 15 and 49 that has experienced some violent victimization at any point in their life (1) or in the last twelve months (2).<sup>9</sup> Taking just a subset of this population and a subset of their offenders – ever married women in this age bracket and their intimate partners -- some organizations estimate that a minimum of sixteen percent and a maximum of 48 percent of such women in most developing countries have been victims of violence.<sup>10</sup> In some of the countries in which DFID works, these percentages correspond to

several million women. If these percentages do not move in the life-time of a country program or within the political cycle of elected governments, it may be tempting to withdraw support, even if individual interventions are having an impact and deliver real benefits to the women and girls they serve.

Both the measure and definitions for these goals may need to be refined. The meaning of the goal to "end" violence is itself ambiguous: it could mean to prevent any new instances of such violence, or it may mean the reduction of its prevalence to zero in a given geographic area. Not only would these two goals require different strategies, or different sequences of strategies, they also would require different indicators. Overall prevalence could decline alongside new instances of violence since violence is more highly concentrated and repetitive in some communities than others.

The chart on the next page helps illustrate this point. According to the results of the Demographic and Health Surveys funded by the US Agency for International Development, 71 percent of the women between the ages of 22 and 44 in Bangladesh report a life-time experience of physical violence by an intimate partner, and 24 percent of these women report a violent incident in the last twelve months. In Haiti and Zimbabwe, by contrast, not only are life-time prevalence rates much lower, but nearly every woman reporting some life-time victimization also reported an incident in the last twelve months, suggesting (if the data are to be believed) that violence is pandemic in certain communities and exceedingly rare in others. Whatever the explanation for these results, a reduction in the number of women reporting a violent incident in the last year in Haiti and Zimbabwe is unlikely to reduce the prevalence rates reported by the population as a whole. Conversely, a reduction in the life-time prevalence of violence among women in Bangladesh is unlikely to reduce the proportion reporting an incident in the last year.

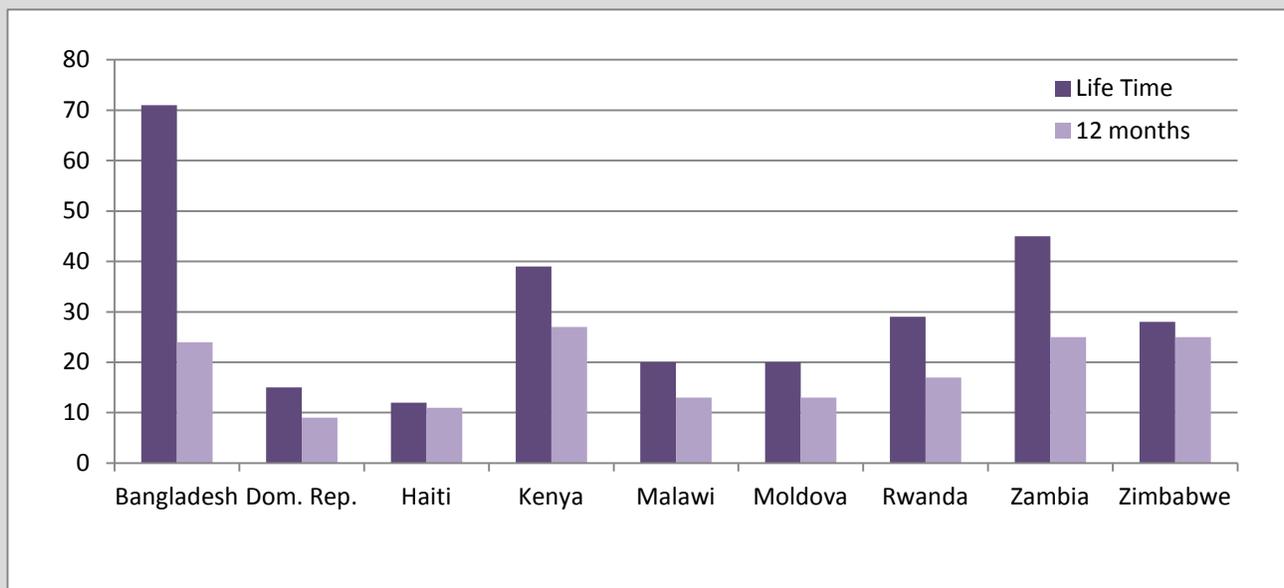
---

<sup>8</sup> The Justice Sector Strategic Plan for Ethiopia, 2011-2015, for example, pledges to increase the rate of conviction of all defendants from 43 to 98 percent.

<sup>9</sup> See the "Report on the Meeting of the Friends of the Chair of the United Nations Statistical Commission on Statistical Indicators on Violence against Women," available at: [http://unstats.un.org/unsd/statcom/statcom\\_2010/Seminars/VAW/default.html](http://unstats.un.org/unsd/statcom/statcom_2010/Seminars/VAW/default.html)

<sup>10</sup> See Sunita Kishor and Kiersten Johnson, "Profiling Domestic Violence: A Multi-Country Study," Calverton, Maryland: ORC Macro 2004.

**Figure 1. Reported Prevalence of Physical Violence, Cohabiting Women Aged 22-44.**



Source: Demographic and Health Surveys, conducted between 2002 and 2006. See Michelle Hindin et. al., "Intimate Partner Violence Among Couples in 10 DHS Countries: Predictors and Health Outcomes," *DHS Analytical Studies 18*, USAID, December 2008, available at: <http://www.measuredhs.com/pubs/pdf/AS18/AS18.pdf>

### Two Responses to the Problem of Scale

Because it is difficult to discern an impact of new programming on the prevalence of violence in the population as a whole, whether it is measured by life-time or annual rates of victimization, some organizations treat the problem of scale as an issue of time. For example, the action plan of the British government's new policy on violence against women and girls reports that "although the goal of the strategy is to end VAWG, we hope that in the short to medium term, there will be an improvement in the reporting of these acts of violence to appropriate agencies." This interim change, it adds, "will yield much higher returns in the future."<sup>11</sup>

As this example shows, all development organizations seek to foreshorten outcome indicators, using proximate signs of improvements in

order to sustain momentum for change. DFID advisors should develop their own indicators for intermediate outcomes, and they should be careful to distinguish indicators of better processes and outcomes, and also to select ones that suit local conditions. Changes in practices of reporting violence, for example, might not indicate changes in the prevalence of the problem.<sup>12</sup> Nor does improved recording of violence presage a future transformation in the quality of the treatment of such violence. Indeed, an increased rate of reporting such violence could easily compromise the ability of service providers, straining capacity or diluting effectiveness. Moreover, this particular strategy and indicator might cause conflict where governments treat reductions in the number of reports of violence as an indicator of good police performance.

<sup>12</sup> See, for example, the recent report of a team of researchers at Harvard and the International Monetary Fund, which attributes an increase in the number of crimes reported by female victims in 17 states in India between 1985 and 2007 to increased political representation of women in government. See Lakshmi Iyer et. al., *The Power of Political Voice: Women's Political Representation and Crime in India*, *Harvard Business School Working Paper 11-092*, July 2011.

<sup>11</sup> The HMG Action Plan, which also insists that VAWG will at some point be "eradicated," can be found here: <http://www.homeoffice.gov.uk/publications/crime/call-end-violence-women-girls/vawg-action-plan>

Another tack is to measure changes in what the World Health Organization and many researchers in the Global North consider “risk factors” associated with violence against women and girls, such as the level of education attained by girls, alcohol use among boyfriends and partners, the age of first marriage, early childhood experiences of violence, and socioeconomic circumstances of key family members.<sup>13</sup> These measures are recognized by most researchers, and some contract implementing organizations organize their work around them in projects to prevent violence and structure services. This approach has deep appeal. It helps resolve the plausibility problem, connecting action now to changes that might be detected soon in a subset of the population with the help of social scientific surveys. It keeps attention on results and the likely impact of development assistance.

Risk factors are proxies for prevalence, though, not a measure of the actual occurrence of violence. Programs premised on risk factors also require special care. Risk factors turn out to be culturally dependent, varying considerably by location.<sup>14</sup> To work well, risk factors have to be locally validated, and they should substantially increase the predicted likelihood of unwanted consequence in order to justify action. Also, risk factors are often used imprecisely in justice systems. As with many insurance schemes, they may over-predict undesired behaviors and thus encourage the treatment of individuals that might not need help or require attention as often as they under-predict real dangers. Because they can trigger unwanted attention to potential victims as well as suspects, programs that rely on risk factors can compromise basic commitments in justice.<sup>15</sup>

More importantly, the use of risk factors circumvents the question of causation, answers to which are

<sup>13</sup> For a careful explanation of some of these factors, see *Preventing Intimate Partner Violence and Sexual Violence Against Women: Taking Action and Generating Evidence*, World Health Organization, 2010.

<sup>14</sup> A recent study from Bangladesh, for example, found that greater women’s “empowerment,” measured by levels of participation in household economic decision-making, was associated with *greater* rates of intimate partner violence. See M. Rahman et. al., “Intimate Partner Violence Against Women: Is Women Empowerment a Reducing Factor?” *Journal of Family Violence*, v. 26 (2011).

<sup>15</sup> For an exhaustive list of risk assessment instruments in use in domestic violence programs in England and Wales, see Sandra Walklate and Gabe Mythen, “Beyond Risk Theory: Experiential Knowledge and ‘Knowing Otherwise,’” *Criminology and Criminal Justice*, v. 11 (2), 2011, 104.

needed in the field. Even with elaborate statistical analyses of data that become available after the completion of research and evaluation, programs that depend on risk factors do not typically uncover mechanisms of activation – why a propensity or predisposition becomes a behavior. Nor do they always clarify the dynamic interactive effects among risk factors and their relationship to “protective factors” that help keep some people safe from violence some of the time. The problem is not the overabundance of variables to analyze, although that aspect of this approach does limit ownership and participation in the development of knowledge to individuals and organizations with specific sets of expertise. The problem is that indicators fashioned along these lines are likely to miss the most important action in development, which typically takes place between the implementation of specific interventions and the outcomes of interest to development organizations. To serve the strategic purposes of development, not just the monitoring and evaluation of individual projects and the achievement of targets, indicators must test a theory of change, generating new knowledge about how activities cause, precipitate, or interact with desired results.

### 3. A Framework for Indicator Development in DFID Programming on VAWG

Below, we outline an alternative framework for indicator development, highlighting three strategies that could help manage problems of scope, scale, and timing in DFID programming on VAWG.

#### Three Step Framework for Indicator Development

- Measure Proximate Changes in the Lives of Women and Girls
- Measure Forces that might Multiply or Minimize the Impacts of Interventions
- Measure the Intrinsic Values of Responding to Violence Against Women and Girls

## 1. Measure Proximate Changes in Women's Lives

Good outcome indicators gauge directly the prevalence of specific kinds of violence in defined communities. They also are sensitive to the kinds of changes that can occur soon as a result of new programs. Indicators of this kind would focus on the experiences of people receiving violence prevention services or help in the aftermath of violence, not the population as a whole or groups with attributes similar to these individuals who do not receive help or new services.<sup>16</sup> For example, you might measure whether victims experience additional violence in the first few days, weeks, or months after participating in a development program, carefully distinguishing the type, frequency, and severity of such violence. You might also measure changes other than violence, such as improvements or deteriorations in the quality of relationships within the family or community, the experience of labor and nutrition and health and education. These and other things may matter greatly to victims' quality of life, even if they are not strongly correlative of current violence or predictive of future violence.

## 2. Focus on Impact Multipliers

Because few programs have enough resources to assist everyone that needs help, and because positive trends can be undermined by counter developments, it is wise to measure changes in factors that might multiply or minimize and otherwise mediate the impacts of interventions DFID supports. These factors include cultural values and attitudes, social norms and habits, and patterns of governance. Some of these things can be found in the attitudes and behaviors of friends and relatives, and also in the views and conduct of government officials. An obvious example is the level of "tolerance" for violence in society, whether it is in families or workplaces. In each society and culture, there may be norms that tend to diminish or discourage violence. They may operate in the same way that women's "empowerment" through employment and other means has been found in some settings to help reduce HIV/AIDs, or in the way that positive deviance and other practices recognized in public health "go viral."

<sup>16</sup> Whether non-recipients of services are equally likely to avoid additional violence is a question about the relative effectiveness of the program, not its contribution to the reduction of violence. The immediate goal should be to ascertain whether women and girls receiving some kind of help experience less violence.

### PRACTICE TIP

#### Impact Multipliers in Programs to Reduce Violence Against Women

DFID advisors might capture the multiplier effects of programs in VAWG by developing indicators about the connections between the kinds of responses to victims of such violence, on the one hand, and women's use of other resources in society, on the other. Measures for these processes would help gauge the empowerment of women.

*"The agenda for the elimination of VAW is not about victimization but rather about the empowerment of women to overcome and eventually change patriarchal hierarchies."*

-- Yakin Erturk, UN Special Rapporteur on Violence Against Women, 2009

Data for these kinds of indicators might be found in other sectors of development, beyond justice and safety programs. Results advisors and researchers in government or academic institutions could help generate answers to relevant questions such as:

- Do women who report violence to the police more often access other state services or receive other kinds of help?
- Do family members or communities in which violence is most prevalent become more or less integrated in development as a result of new programs?

These impact multipliers might be poorly understood and yet still be measureable. That is, one might develop an indicator of the transmission of the norm behind an intervention without knowing precisely the means by which it grows, disperses, or catches on. Take, for example, a training program for para professionals that first encounter victims of sexual assault in a clinic. Improved training might be critical for the effectiveness of their work but it is unlikely by itself to directly benefit women and girls with whom they have no direct contact. So, one might use the occasion of a public announcement of the training program to detect whether other members of the community notice, appreciate, and share the commitment to counter such violence. One need not ask if the announcement itself or the training has an independent impact on community norms. It would be enough to know if members of the community observed, believed, and welcomed the new commitment. The recognition of others that

something new is being done on VAWG would indicate meaningful impact, and perhaps receptivity.

### **3. Measure Intrinsic Values**

Another strategy is to appraise changes in experiences of justice that have intrinsic value, ones considered worthy and important regardless of their contribution to the amount of violence in society. Whether it is the sense of being heard, taken seriously, or treated professionally and with dignity, some experiences may be especially worth supporting and developing, even when improvements can only be incremental and hard to measure. For example, in some places, whether or not victims are “satisfied” by the operation of justice system may have particular importance, especially if in the past the system was widely seen as dismissive or insensitive to people’s needs. In other places, an easing of the burden of asking or receiving help may have special significance for victims, regardless of whether this is a harbinger of their likely recourse to the justice system upon the experience of any problems in the future. The values that matter intrinsically are likely to vary across countries and conditions.<sup>17</sup>

The three strategies will work best in combination. All of these strategies would likely require local surveys at the outset in order to measure the prevalence of the problem among groups targeted in the program. Focus groups with victims as well as non-victims might help identify coping strategies and other things that alleviate the burden of violence. They might also identify the intervals at which victims experience further violence, helping you to tailor the timing of measurements as well as interventions.

### **4. Development Scenarios, Theories of Change, and Justice Indicators**

To make these principles concrete, we describe below two different development scenarios. These examples should illustrate how to convert an opportunity into an indicator by articulating one or

more theories of change and then proposing a set of measures that would test them.

Imagine you are trying to reduce violence against women and girls in a country in which the ministries of education and culture are committed to introducing new instruction in secondary school that teaches boys respect for girls. Meanwhile, the ministries of the interior and justice as well as the national planning institute are dismissive of the plan and skeptical about their likely impact on levels of violence. Imagine, further, that there is no evidence that school-based instruction will have this effect in this particular country, although similar undertakings are reported to have this consequence elsewhere. Assuming the support of your country program and regional results advisor, what kinds of indicators would you develop to guide the implementation of the project and report results?

First define some of the ways in which teaching boys and girls to interact in more respectful and peaceful ways in the course of their education might reduce violence against women and girls. One way is by lowering the level of violence during or after school among participants in the program. Another possibility is that these lessons continue to operate as participants mature, shaping their conduct as adults, too. A third possibility is that the behaviors and norms acquired during such instruction somehow infects others as well, permeating the school’s culture or even the larger community it inhabits, and possibly suppressing some of the countervailing forces in society.

If this was the theory of change, you would benefit from at least three indicators: one that captures change in the amount of violence reported among and by children in school, distinguishing participants from others; another that tracks the longevity or stickiness of the learning; and a third that captures changes in patterns of violence in the surrounding setting. You might also need a measure of the extent to which the project has perverse effects, since parents who learn about reports of school violence might themselves become more violent, either in respect of their own children or others.

Or take another scenario. The national policing organization of a country, backed by its supervising ministry, is keen to expand a new form of “community policing” that, according to an international assessment team, increased the portion of the public that has confidence in the police in one province. The commissioner of the police encourages you to support the extension of this

---

<sup>17</sup> It is customary in the Global North to inquire about the level of victims’ “confidence” in the justice system, but primarily because of beliefs that public confidence improves the effectiveness of criminal justice, not because trust, hope, and confidence are deemed independently worthy. See, for example, Ben Bradford, “Voice, Neutrality, and Respect: Use of Victim Support Services, Procedural Fairness, and Confidence in the Criminal Justice System,” *Criminology and Criminal Justice*, 11 (4), 2011.

**Figure 2. Possible Indicator Structure for Two Development Scenarios**

	<b>Proximate Measure of Change</b>	<b>Potential Multiplier</b>	<b>Intrinsic Value</b>
<b>Scenario 1 (school-based instruction for boys on respect for girls)</b>	Change in number or frequency of reports of violence among program participants a short time after completion (e.g., 1 month)	Change in number of students not participating at the school who approve of the program and consider it worthwhile	Proportion of participants that feel/believe they are better able to protect selves from serious harm
<b>Scenario 2 (greater access to justice and community policing)</b>	Increase in number of reports of <i>minor</i> crimes or violence against women or young children	Proportion of women accessing justice services who have heard about others' good experiences	Proportion of women who say they appreciated and benefited from contact with the police

particular scheme to a second province, claiming that it will also reduce violence against women and children. Again assuming the support of your country program director and results advisor in the policy division, what kinds of indicators would you want in order to manage the project and assess the outcomes?

Start by defining why a program to expand public access to a community policing scheme might lead to less violence against women. One possible theory is that women who in the past did not turn to the police after experiencing some form of violence would now do so, and that they would gain something in the course of that interaction that reduced the amount of violence they personally experience in the near future, over and above any reduction that might occur on its own. For example, they might receive help from the police that would help them avoid future victimization, or those responsible for the violence would receive some restorative or punitive consequence that had this effect, or both. These same women might not only turn to the police again in the event of any future violence, however severe, but they would also communicate to other women their positive experiences, leading other women and perhaps others as well to increase their use of police services.<sup>18</sup>

If this, then, was the theory of change, at least two indicators would be helpful: 1) the proportion of women who are return complainants, seeking help from the police for any problem, whether it is theft of

property, a need for help with a neighbor, or another experience of violence; 2) the percent of women who have heard of other women's good experiences with the police or told others about it.

A second theory is possible in this scenario, too, and may be equally worth testing. For example, through some combination of professional development and new management in policing, the police might become a civilizing, calming, empowering public service, whose educative and preventative impact on the level of violence in society is positive. In this theory, instead of the response to individual victims being the operational key, it is the active engagement of the police with the public that causes the reduction in violence. This theory would call for a different set of indicators, such as the percentage of the public that recognizes that the police are a source of support for women suffering from violence.

The differences in these theories of change matter greatly, as do the implications of the indicators attached to them. If it is the actual interaction between individual women and police that reduces violence (whether it is via help, sympathy, or accountability), then further progress as well as any expansion of the effects to other provinces will depend on a resource intensive increase in retail work in the justice sector. If on the other hand it is the impersonal civilizing norms that suppress violence, then there might be other ways of achieving the same ends.

Notice that for both scenarios we have picked a proximate measure of change, proposed an indicator for an impact multiplier of the intervention, and emphasized a change that might appeal broadly to the values and interests of the public irrespective of any changes in prevalence. In the first scenario, as figure 2 above shows, the expected change might be a near-term reduction in incidents of violence among

<sup>18</sup> Even if the combined effects on levels of violence hypothesized here were negligible, or diminished or even negated by countervailing factors such as a major upsurge in violence against women or the mistreatment of victims by police in several areas, the theory of change might still be correct. A measure of the prevalence of violence against women and girls in society alone would thus miss the important movement here, overlooking changes in key relationships on which other programs might focus future interventions.

**Figure 3. Indicators Relating to VAWG in DFID Log Frames in Four Countries**

Occupied Palestinian Territories	Nepal	India
% of public in the West Bank who are at least satisfied with security and justice services.	% of women and children in VDCs where PLCs have been established, who feel more secure and protected from violence and abuse than before the PLC was created.	A 2.5 million reduction in women reporting violence in the last year across the three states.
% of public who identify improvements in accountability of security services.	% of women and children in VDCs where PLCs have been established, that are more willing to report violations and have more confidence in their ability to access justice and support than before the PLC was created.	
Specialist justice and security sectors professionals and systems investigate and prosecute cases of domestic violence.	Number of women and children reporting violations to the PLCs and the % who are satisfied with the outcome, disaggregated for excluded groups.	

*Note: this is a partial list of indicators used in the country programs.*

program participants. The possibility that norms of respect learned in the program might migrate and improve the lives of other students could be tested by a measure of supportive attitudes among non-participants. And by measuring the proportion of girls in the program that feel capable of avoiding serious harm, one could reinforce the value an educational experience without violence.

Figure 2 also contains possible indicators for the same dynamics in the second scenario, above, focusing on a proximate measure of change, a potential multiplier of its effects, and an intrinsic value of a more respectful and accessible community policing scheme.

The main lesson here is that good indicators interrogate theories of change. They must coax out the premises lurking beneath inexplicit or vague theories of change, and they must drive attention to key propositions. Indicators must also be developed for negative possible outcomes – the possibility that the preferred theories of change may be wrong and that other forces and factors might be more important in reducing (or aggravating) violence than the ones advanced in the specific project. In this way, indicators can facilitate learning about outcomes that are expected as well as those that are not.

## 5. Indicators Currently Used in DFID Programming

We were only able to review a few of the indicators in three country programs, so we cannot tell whether the ones described below are representative of the measures in all programming. The indicators we studied, moreover, appeared in log-frames, which do not always fully clarify relationships between the expected outcomes and the array of activities being supported and described in greater detail in program documents. Below, then, we explore theories of change that might support these indicators, and review the kind of information that would reliably generate their measures.

As figure 3 above shows, the indicators used in the occupied Palestinian territories, Nepal, and India all focus on the expected results or impact of interventions, although the outcomes are sought at different levels of society. In the OPT and India, for example, the impact sought is at the level of the population, whether it is the public as a whole (OPT) or all women (India). In Nepal, the impact sought is in the communities in which interventions are taking place, in the village district centers (VDCs) where paralegal committees (PLCs) are being established and among women and girls. Only in India is the outcome a measure of the prevalence of violence against women.

We heard only brief accounts of the histories of these indicators from advisors, some of whose programs were inherited or still under design, and we read only short descriptions of the activities being funded in each location. We recognize, moreover, that some of these indicators were developed under pressure – “in a bit of a rush,” according to one advisor – and in the absence of reliable, sufficient, or ideal information. Our goal in this section is not to question the logic behind the design of the program, but rather to highlight stages and processes of indicator development that could generate knowledge about how these expected results might come about.

### *Occupied Palestinian Territories*

According to the log frame, the main purpose of the program in the OPT is to help establish a justice sector that is “more accountable and responsive to the public.” The activities through which this might be achieved include specialist training for justice system personnel that respond to problems of violence against women and girls, the creation of family protection units within the police, the operation of shelters for battered women, more and better counseling in custodial rehabilitation centers, and improvements in the operation of a complaints system in the security sector.

The main outcome indicator – “the percent of the public in the West Bank that are at least satisfied with security and justice services” – is quite distant from any conceivable measures of the potential benefits and effects of shelters, counseling services, and the responses of family protection units in the police. Even with reliable survey-based data on public perceptions, disaggregated by sex, it would be hard for anyone to understand how even a well-integrated set of small-scale interventions, designed and operated for small numbers of women and girls, might have an impact on the views of the population as a whole. Moreover, these public perceptions could move in any direction as a result of developments wholly unrelated to program activities, or even events within the OPT.

So, what might you do? Would you throw out the results indicator? Or could a theory of change buttress it? The second indicator in the OPT log frame might help, since it contemplates a multiplier, suggesting that the percent of the public that “identify improvements in accountability of security services” might increase over time. And yet that indicator also might move in ways unrelated to program activities, as a result of media reports,

foreign policy events, and other developments. So one task would be to specify possible ways in which members of the public not directly experiencing violence against women and girls or witnessing the operation of the justice system could notice or believe things are improving. An indicator that gauges interactions between recipients of services at shelters and family protection units and others who do not have such experiences might help, especially if it registers their perceptions of the value of these services. If friends and relatives of victims notice improvements in the responsiveness or accountability of security services, this would strengthen the measure of the intrinsic value of responding to victims of violence as well as the likelihood that this response matters to the public as a whole. At least, such a measure would gauge the cultural motility of improved services and more serious state responses to female victims of violence.

What about credible measures of proximate changes? Here the challenge is less of a technical than political nature. In Palestine, the government reportedly sees the new “family protection units” principally as a means of advancing the cohesion of families, not just the interests of injured women and girls. The contract implementing organization, UNWOMEN, struggles to collaborate productively with the Ministry of Interior, whose interest in the speedy “reintegration” into families of female victims of violence is sometimes directly at odds with the organization’s commitment to the physical protection of battered women. While the government and UNWOMEN might find separate ways to measure what they most care about, DFID needs an indicator of the contribution of the FPUS to the safety of women and girls in society, and it relies on other entities for the data and knowledge on which any such indicators would depend. It also needs an indicator that the government won’t quickly orphan.

One possibility is to canvas the list of more than 100 indicators proposed in the “strategic plan for combating violence against women,” prepared by the Ministry of Women’s affairs. And yet, even if an indicator in that document matched DFID’s purposes, it might not have traction with the organizations directly responsible for the administration of the new programs for women, whose professional interests would strengthen the likelihood of the use of a locally crafted measure. In these circumstances, it might be wise to build indicators on the basis of what staff in the FPU would consider a sign of the reduction of violence

among women who have sought their help. This could be the number and severity of repeat incidents, the total amount of violence in the families, homes, and relationships from which the majority of incidents are reported, or some other measure entirely.

### *India*

The outcome indicator in India focuses squarely on prevalence, projecting a 4 percent reduction in the total number of women aged 15-49 reporting violence in the last 12 months across four states, each with relatively high rates of reported victimization according to the last national survey. The theory is that a wide array of new interventions (help hot lines, short stay homes, task forces, trainings, health services, alcohol management groups, batterer programs, legal services, martial arts training) will contribute to the reduction. The program resolves the problem of scale by proposing to measure rates of violence in the 30 districts in which new programs are operating. It projects a 10-30 percent reduction in the prevalence of such violence in these areas, which, added together, would represent a 4 percent reduction in the incidence of violence for the entire female population in three states.

Even with the help of this math, observers of the innovations in India might need additional tools in order to appreciate whether individual interventions work and how they deliver dividends directly to women. Using the hotline as an example, we can outline four hypotheses for how it might contribute to this end. One possibility is that women who call the line might be exercising a form of self-help, and that this act by itself, even without anyone picking up the phone on the other end, could contribute to reduced incidence of violence in their own lives. Another possibility is that men might be deterred from acts of violence as a result of knowing about the existence of the hot line, whether or not it is used. A third possibility is that the response to some or even all calls might affect the future conduct of the victim or perpetrator in the relationship that generated violence.<sup>19</sup> A final possibility is that the hotline itself would contribute to a cultural message that violence

should not be tolerated, and that this message is received and helps depress the incidence of violence on its own while reinforcing and multiplying the effects of the first three hypothesized mechanisms. An indicator for each of these possibilities would generate the kind of knowledge and evidence that governments, civil society actors, and development agencies in India and other countries need to try out their own ideas about reducing VAWG.

A final challenge here, as program documents make clear, is to understand how the interventions work “in combination” with one another, and also how more generalized “awareness raising” and “community attributes related to violence” in particular are related to the incidence and prevalence of VAWG. Some of this knowledge can be gained through complex statistical analyses of the variables involved in the interventions, all of which will be facilitated by the use of randomized control trials. Yet few people in the area are likely to have the skills necessary to participate in or appreciate the results of this kind of inquiry. To expose the multiplier, the program might survey the population and ask residents whether they notice and believe the priorities of the government of the states in which the interventions are taking place are changing. Changes in these perceptions might not result from or be attributable to DFID-based activities, but that is less important to know than it is to ascertain whether or not some potential multiplier of DFID’s support for reducing VAWG is active in the community.

### *Nepal*

The indicators in the Nepal program focus on the experiences and perceptions of women and girls in locations where the program’s most visible innovation – paralegal committees (PLCs) – have been established. Although there is no direct measure of the impact of these new services on the experience or prevalence of violence, there appears to be an indicator of the intrinsic value of these services – the proportion of women who are “satisfied with the outcome.” The program also assumes a multiplier or cultural meme since it projects an increase in the proportion of women and children in VDCs where PLCs have been established that are “more willing to report violations,” and “feel more secure and protected,” although it doesn’t specify how this would take place and propose a means of measuring it.

---

<sup>19</sup> The paper has not discussed indicators of changes among suspects and offenders. We found little information in current programming that expressly addresses this dimension of efforts to reduce violence against women, which is in contrast to the priorities in many wealthy countries, where work with batterers and other kinds of offenders are a familiar part of multi-modal strategies to reduce violence.

Both indicators would be stronger with an explicit theory of the processes by which beneficial effects of the PLCs will circulate in society. Program documents imply that many such mechanisms are in place – educational events, dissemination campaigns, cultural celebrations – and also they may be as or even more important than the actual services. The reputation and effects of these kinds of interventions are equally important to measure, even if they might appear as frippery. Observers of the program in Nepal also will need to know whether the mere presence of a new support system provides reassurance to women or whether the quality of their experiences with that system matters more.

There also needs to be a measure of some proximate change in the experience of women attended by the PLCs. Both the indicator and theory for this indicator should draw on what the PLCs do for women, whether they connect them to social services, furnish representation or assistance in legal matters, or provide some other form of help. Without knowing more about the design of the PLCs, speculation about the best measures might impose an outcome or effect that matters elsewhere more than in Nepal. A useful measure of impact might result from conversations with the women receiving services.

## **6. Collecting, Managing and Evaluating Data on VAWG**

DFID advisors face multiple challenges in the collection, management, and evaluation of data that relate to violence against women and girls. There are, for starters, no truly global estimates of the prevalence of physical and sexual violence against women and girls today. Multi-national surveys are conducted irregularly in a small number of countries and generate measures primarily for levels of intimate partner violence, which may not be the central concern in every development context. The evidence available to design interventions in any single country is also slim. Few programs are proven to work in multiple locations. Most scholars are circumspect about the kinds of claims about causes and effects that can be made responsibly.<sup>20</sup>

<sup>20</sup> A team of researchers from the London School of Hygiene and Tropical Medicine recently concluded: “although there is no magic bullet to reduce partner violence, the consistency of our findings across sites suggests that a prevention strategy, once validated and refined, might have relevance in a wide range of settings.” See Abramsky et al., “What Factors Are Associated

Indicators of the consequences of VAWG are generally stronger than measures of their causes, and the latter vary considerably by country and community.<sup>21</sup>

The ethical and methodological challenges in the collection of appropriate data and knowledge are acute. Researchers must avoid causing harm to women and girls in the course of learning about the scale and sources of violence they experience. Typically this requires more than earning the loyalty and confidence of women in vulnerable circumstances, a long-term undertaking in its own right, even for organizations that directly provide services to victims. Fortunately, there are excellent guides to these challenges that can help advisors meet the obligations of development practitioners.<sup>22</sup>

DFID advisors can help by participating in the collection, management, and evaluation of data for the indicators in need of development. The data required for indicators in DFID programs do not exist in nature: they must be manufactured or assembled from imperfect sources and then interpreted. There are limitations to all sources of these data, whether they come from the records of non-government organizations providing services to women and children, the routines of government departments, or responses to questions in population surveys. Here we describe three special challenges, suggesting a ways DFID programs could improve the empirical foundations on which indicators are built.

### *1. Strengthening Survey-Based Measures of Violence Against Women and Girls*

DFID advisors could help improve upon the most common source of estimates of the prevalence of VAWG – a survey of a representative sample of the population – by ensuring that they detect changes over periods of time that are aligned with programs,

---

with Recent Intimate Partner Violence,” *BMC Public Health*, 2011, 11: 109.

<sup>21</sup> See, for example, “The Effect of Intimate Partner Violence and Other Forms of Violence Against Women on Health,” *Journal of Public Health*, v. 33, no. 1 (2011). Compare also the findings in Abbi Kedir and Lul Admasachew, “Violence Against Women in Ethiopia,” *Gender, Place, and Culture*, v. 17, no. 4 (August, 2010), and Amber Peterman et. al., “Estimates and Determinants of Sexual Violence Against Women in the Democratic Republic of Congo,” *American Journal of Public Health*, v. 101, no. 6 (June 2011).

<sup>22</sup> See, for example, Mary Ellsberg and Lori Heise, *Researching Violence Against Women: A Practical Guide For Researchers and Activists*, WHO and PATH, 2005.

policies, and governments. This means that surveys must be administered on a recurring basis, with funding for at least five years since changes in the design are inevitable in response to early discoveries, and programs need at least three consecutive years of data that come from the same questions.<sup>23</sup> Surveys also must also inquire about onset, frequency, and severity of violence, including the time that elapses between critical incidents as well as about its desistance. In order to generate insights about how to intervene in cycles of violence, surveys must investigate the experience of repeat victimization and the context in which it occurs.<sup>24</sup> Finally surveys must ask about the experience of help, whether or not it comes from a state institution or other recognizable source. Surveys must teach you about how friends respond to victims of violence, and the coping strategies women develop in the absence of formal systems of assistance.

Because prevalence rates are unlikely to move much or quickly, DFID programs should consider other measures of outcomes, some of which might come from surveys. For example, one potential measure is the proportion of women, both victims and non-victims, that believes there is nowhere to turn for help in the event of violence. Another potential measure might come from victims' evaluation of the government's services, whether women thought they were helpful, taken seriously, or otherwise found value in the response.<sup>25</sup>

A third improvement would be to use the results of surveys more often to answer other people's questions about justice and safety. The results of many surveys are used at only one point in time for the purposes of advocacy. Where, though, the results of surveys are mined over time and create insights about theories of change in this sector, whether or not these are directly related to VAWG, they may

become common resources for public policy change and gain traction.

## *2. Improving the Use of Administrative Data*

The results of surveys will be stronger when combined with administrative data, no matter how frail each of these sources may be. One reason is that the leaders of government agencies, including those in the justice sector, do not always recognize or acknowledge the measures of the prevalence of violence against women and girls generated by internationally designed and implemented surveys. Also, national surveys rarely capture local patterns and problems in ways that make sense to sub-national leaders. As a result, according to researchers at the Geneva Declaration, "local officials often do not have sufficient information about the scale of VAW in their communities."<sup>26</sup> Local officials also may not see themselves and their work reflected in the indicators, leaving them unmotivated to move the measures. To encourage local officials to respond to estimates of the problem with VAWG, indicators may have to more closely align with the systems of authority in local administrative bodies.

One way DFID advisors could make better use of administrative data for the purpose of understanding the scale or sources of the problem with VAWG is by using police data on recorded crime. There are plenty of good reasons not to use this source, including the partial nature of reporting and registering these problems and the incomplete manner in which information is collected. And yet, if used in combination with statistics from other sources, such as hospitals, registries, and non-governmental organizations providing services to women, police data can help paint a portrait of the problems with VAWG that law enforcement agencies recognize, and to which they may be held accountable. By taking this data seriously, moreover, DFID advisors can help add demand for its improvement.

Almost all law enforcement agencies around the world record the sex of the victim, a practice that might permit an understanding of the relative burden of crime on women in any society, no matter

---

<sup>23</sup> DFID may need to find ways to reduce the relative costs of surveys. Some DFID programs have found ways to fold them into national projects of the study of public health, which effectively subsidizes their costs and also makes them more likely to be reproduced on a regular basis.

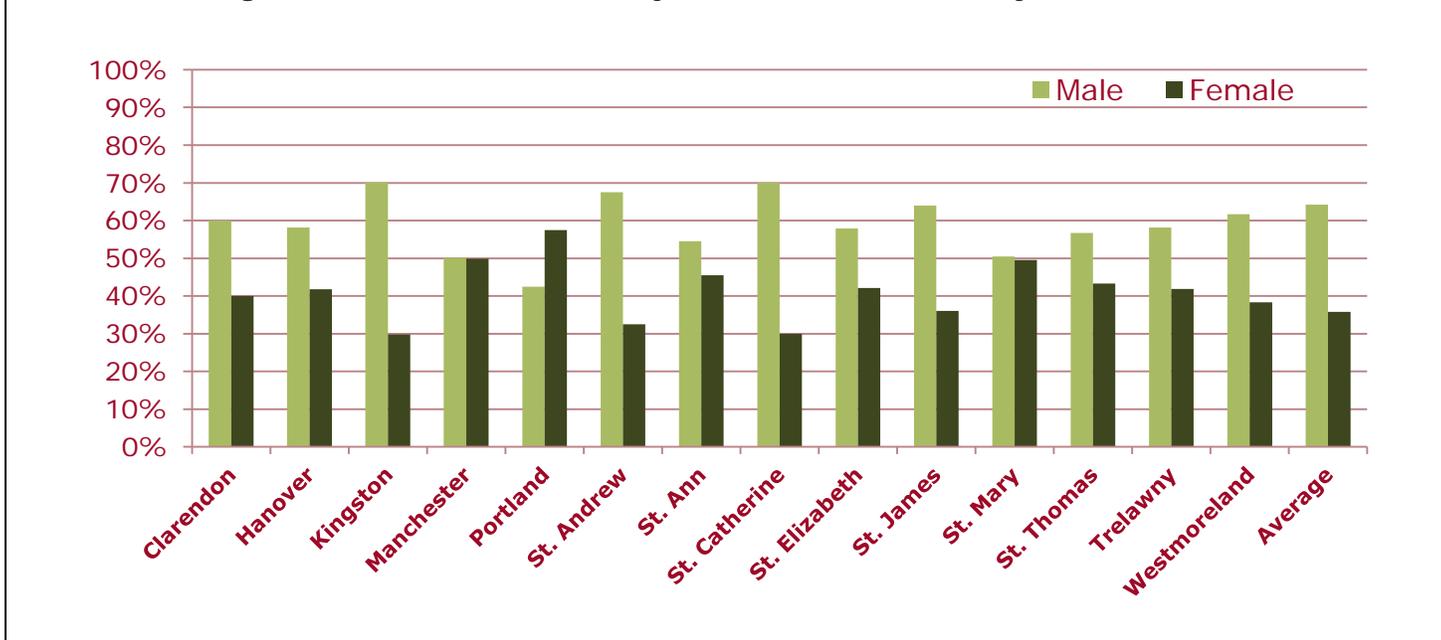
<sup>24</sup> The International Violence Against Women Survey, which in 2010 included two developing countries – Mozambique and the Philippines -- is one of the few surveys that generates findings about repeat victimization. See the reports from HEUNI, available at: <http://www.heuni.fi/12859.htm>

<sup>25</sup> In many countries, victims' evaluations of the justice system are negative, so there may be a low starting point, and it may be helpful to distinguish between the evaluations of victims of grave or repeat violence.

---

<sup>26</sup> Elisabeth Gilgen and Jesna Lazarevic, "How to Systematically Measure and Monitor Violence Against Women," Background paper, Expert Workshop on VAW – Disabling Development," March 2010, Geneva Declaration, <http://www.genevadeclaration.org/en/measurability/monitoring-armed-violence/indicators-of-violence-against-women.html>

Figure 4. Sex of Victims of Major Offenses Recorded by the JCF, 2005



how incomplete the data may be. Using data from the Jamaican Constabulary Force, the chart above shows that women comprised half of all victims of major offenses in two divisions (Manchester and St. Mary), compared to the national average of one third. The finding is not at all diagnostic, and it could not be used as an indicator of the prevalence of VAWG since it would move in response to increases in crimes that predominantly involve male victims (such as shootings). But the finding could be used to cue up a conversation with the government about the best distribution of resources to prevent and respond to crimes involving female victims. In contexts where surveys are not available, where policy conversations take place without reference to empirical information, or where the government is disinterested in new measures of the problem with VAWG, police data may be a place to start.

### 3. Defining Norms and Standards

Many of the norms operating in the field of VAWG today come from outside of developing countries, including legal and statistical definitions of violence, juridical principles for the evaluation of evidence, and basic rules in criminal and civil proceedings. Even notions of what constitutes a healthy rate of reporting violence to state officials, or what proportion of prosecutions should end in a conviction, tend to come from abroad. Many government officials and non-governmental leaders labor under the assumption of the superiority of the

rates reported in the Global North, and use them implicitly or explicitly as a standard by which to judge local practices. Some civil society organizations complain that “only” 28 percent of rapes are reported to the police, and that “only” 50 percent of victims are satisfied with the response of the police. This is a world-wide concern, but there is no clear norm or standard which helps discipline the interpretation of the rates recorded in any one country.

DFID advisors could help national leaders in developing countries contribute to the creation of norms for this field by helping them to review some of the assumptions behind such indicators. Some of the suppositions about average rates of reporting and victim satisfaction need to be corrected, as might the scale of promised improvements. Where national governments propose giant changes in short periods of time, such as a doubling of a rate of conviction, DFID advisors might temper claims and recommend incremental yet steady improvements over time. A critical review of practices in the Global North would find that indicators for police and prosecution and courts are rarely fixed and often are refined in response to new policies and governments.<sup>27</sup> Candor

<sup>27</sup> The Crown Prosecution Service for England and Wales, for example, has yet to finalize the indicators it will use for evaluating progress in implementing the government’s action plan against VAWG. See Annex D in the CPS action plan, [http://www.cps.gov.uk/publications/equality/vaw/vaw\\_strategy.html](http://www.cps.gov.uk/publications/equality/vaw/vaw_strategy.html)

about the status and evolution of indicators of VAWG in the Global North might help national governments in developing countries see themselves as partners in a truly global enterprise.

## **Conclusion**

There may be no universal, optimal, or even just consistently superior strategy for measuring the prevalence of violence against women and girls, the effectiveness of the response to it, and the evaluation of the impact of interventions designed to reduce one or another form of VAWG. This is in part because the field is so young. The systematic study of violence against women by international organizations began less than two decades ago, and it may not have started in all of the countries of DFID programming. Another reason is that the forms and etiologies of violence against women and girls are too diverse for there to be a standard set of best practices: efforts to reduce fatal forms of VAWG require different baseline research, impact assessment tools, and indicators of progress than programs to diminish workplace battery, domestic abuse, stalking, and the sexual assault of young girls. A third reason is that the effectiveness of any development program depends primarily on its fit with the political environments in which efforts to improve the lives of women and girls are implemented, and these environments diverge widely around the world, varying greatly even within individual countries.

To capture the imagination of government, harness the ambition of officials, and reward the efforts of front-line workers, measures and indicators have to make sense locally. This paper has suggested a framework in which DFID advisors might facilitate local development of indicators, starting with choices about strategies and indicators for reducing and responding to violence against women given the gap between the scale of the problem and the means available to reduce it. It outlined a process by which indicators might be constructed, starting with the joint definition of a problem, its contextualization, and articulation of a theory of change, followed by a hunt for relevant data among available systems of information that could routinely generate the measure. In practice that process is more zigzagged than linear, and, like most development work, requires diplomacy. It will also need to be replenished over time. The ending of violence against women is a long-term undertaking. Were it to be close at hand, the field would not need a rich array of indicators.

*Todd Foglesong is a Senior Research Associate at the Program in Criminal Justice Policy and Management, John F. Kennedy School of Government, Harvard University. He would like to thank DFID advisors in many countries for their helpful comments and insights.*

**Program in Criminal Justice Policy and Management**

Harvard Kennedy School  
79 JFK Street  
Cambridge, MA 02138

+1 617 495 5188

[www.hks.harvard.edu/criminaljustice](http://www.hks.harvard.edu/criminaljustice)