GUIDELINES FOR IMPACT OR OUTCOME EVALUATION

FOR PROJECTS FUNDED BY THE UNIFEM TRUST FUND TO ELIMINATE VIOLENCE AGAINST WOMEN

Produced for the Gender and Development Group – PREM
World Bank

by Linde Rachel
April 2006

CONTENTS

ACRONYMS

CDC	Center for Disease Control
CDD	Community Driven Development
CEDAW	Convention on the Elimination of Discrimination Against Women
CIS	Commonwealth of Independent States
CSPRO	Census and Survey Processing System
DHS	Demographic and Health Surveys
GBV	Gender Based Violence
IDRC	International Development Research Centre
IEG	Independent Evaluation Group
LFA	Logical Framework Approach
LSMS	Living Standards Measurement Study
MICS	Multiple Indicator Cluster Survey
NGO	Non Governmental Organizations
RRA	Rapid Rural Assessment
SIF	Social Investment Fund
SMART	Specific, Measurable, Achievable, Realistic, Time-related
SPSS	Statistical Package for the Social Sciences
TF	Trust Fund (UNIFEM)
UNDP	United Nations Development Programme
UNIFEM	United Nations Development Fund for Women
VAW	Violence Against Women
WB	World Bank
WNGO	Women's Non Governmental Organizations

INTRODUCTION

Background

The UN Trust Fund in Support of Actions to Eliminate Violence Against Women was established by General Assembly Resolution 50/166 in 1996 and is managed by the United Nations Development Fund for Women (UNIFEM). In establishing the Trust Fund, the General Assembly highlighted eliminating violence against women as critical to accelerating the implementation of the 1995 Beijing Declaration and Platform for Action. Acts or threats of violence against women, wherever they occur, constitute a violation of human rights and an obstacle to the achievement of equality, development and peace.[1]

UNIFEM began its work in the area of violence against women in 1993, when pressure by civil society organizations on their governments had led to the UN Declaration to End Violence against Women, and in accordance with its support to women's networks organizing for the World Conference on Human Rights in 1994. UNIFEM was one of the first UN agencies that translated the UN Declaration of 93 and the Platform for Action of 95 into concrete programs that aim to end VAW. [2]

According to the Fund's operational guidelines, the emphasis of the trust fund is on "innovative models and strategic interventions", providing funding for innovative and catalytic initiatives that seek to prevent and eliminate violence against women.[3] "The idea is to provide grants to projects that can later serve as examples for replication in other countries." [4]

The Trust Fund invites proposals from civil society organizations worldwide that have experience in local gender-related issues. It disburses approximately $1 million a year, which funds only a small percentage of proposals received (for example, 18 out of 270, or 7%, in 2002). Through 2004, individual grants have been relatively small, averaging about $50,000 for periods of 1 to 2 years. Recently, the Trust Fund committed itself to making larger and longer grants, which will range from $100,000 to $200,000 for periods of 2 to 3 years.

Project proposals are selected and funds are granted on an annual basis by a global Program Appraisal Committee (PAC) composed of at least 9 UN organizations and a small number of international NGOs working on ending violence against women. In 2005, the Trust Fund decision-making process was devolved to regional review committees in an effort to bring decision-making closer to the ground and secure greater involvement of the UN Country Teams in deciding on, monitoring, and linking with Trust Fund programs.

Since its inception, the trust fund provides not only for project implementation, but also for the evaluation of project impact. Project proposals generally set aside 10% for project evaluation, which corresponds to the general rule of thumb. In recent years the importance of impact evaluation to the trust fund management has been underscored by separate guidelines for

[1] United Nations, General Assembly: A/RES/50/166 (99th plenary, 22 December 1995), with reference to the Vienna Declaration of 1993.
[2] Desk Review of UNIFEM's Work to End Violence Against Women, submitted by the Education Development Center. 14 April 2002,
[3] Operational Guidelines for the Trust Fund in Support of Actions to Eliminate Violence Against Women, 2001.
[4] Introductory statement from the UNIFEM Trust Fund website.

project impact assessment that are attached to proposal guidelines. In practice, however, the projects are rarely subjected to systematic and technically defensible impact evaluations.

At a workshop held by the World Bank on the Development Implications of Gender-Based Violence in November of 2004, workshop participants, including UNIFEM representatives, identified the need to conduct systematic evaluations of the trust fund projects in order to determine their effectiveness in addressing gender-based violence.[5] The associated recommendation was to develop an evaluation guide that UNIFEM can use to conduct impact evaluations of its existing and future portfolios of projects.

The current document represents the first step in this endeavor: the drafting of a guide to impact evaluation based on a representative sample of fifteen projects, drawn randomly from the list of Trust Fund projects funded and completed to date.

Selection of Trust Fund Projects

Using the complete listing of projects that had been completed to date, a systematic random sample of fifteen projects was selected.[6] Six of the projects had been funded prior to the year 2000 and nine between 2000 and 2004. The majority of the grants (9) received funds between $ 40-60,000. Five projects received less and only one received more ($110,000). The duration of the funding period was one year for 12 out of the 15 projects. Three projects exceeded that duration to a maximum of 22 months. A more detailed description of each project is provided in Appendix 1.

From its inception until 2004, the guidelines for grant-making decisions described strategies to be supported under six general headings: Awareness raising and advocacy, capacity building for organizations, legal literacy projects, gender-sensitivity training, violence prevention initiatives, action-related research. These are clearly not discrete or mutually exclusive categories. All projects included some element of awareness raising and/or knowledge increase in their strategic objectives, and for most of them (13) these are combined with other change objectives. In the list below, capacity building is implied in most of the objectives aside from empowerment.

Increased awareness and/or knowledge – only	2
Increased awareness and/or knowledge – plus	13
Advocacy / policy change	11
Empowerment	3
Skill development	3

The overall profile of the selected projects reflects a considerable need for groundwork at the level of awareness, knowledge, attitudes and practice in complex social and cultural environments. It also portrays the human rights approach of UNIFEM that stresses ground-up activism.

[5] The primary objective of this workshop was "to examine the negative impact that violence against women has on development and to explore how we could better incorporate issues of gender-based violence into our operational work and promote a greater focus on such issues on the part of our client countries." (Ingram, 2005).
[6] That is, projects that had received the final 10% of the grant after their final evaluation report.

Since the focus of impact evaluation is on outcomes rather than process, we reorganize the projects by the following key outcome categories:

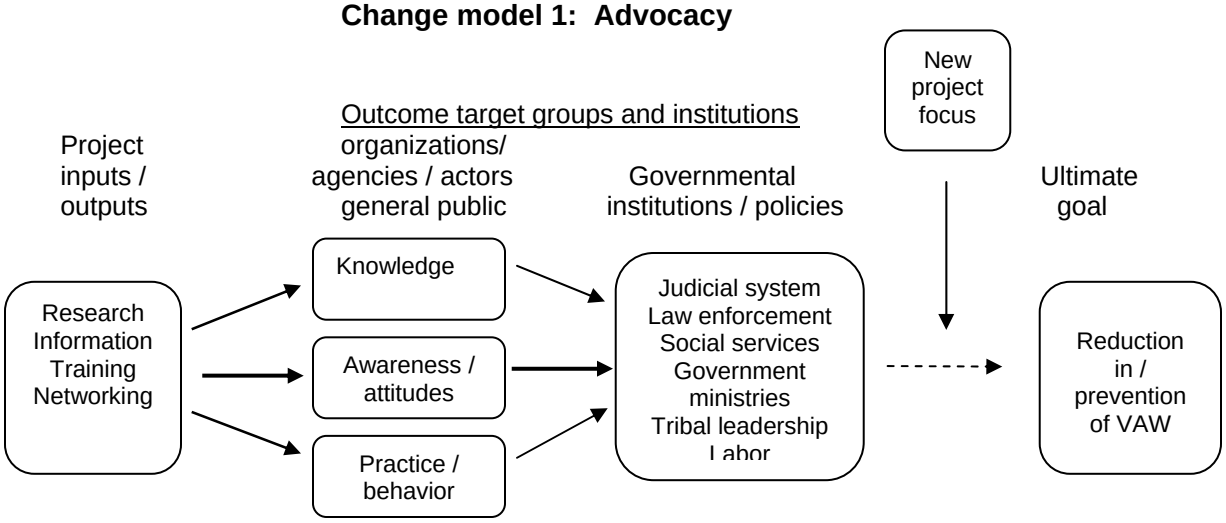
A Increased awareness / knowledge
B Increased advocacy-related outcomes[7]
C Increased capacity (legal literacy, empowerment, service provision)

It should be kept in mind, however, that in the sampled projects, these outcome categories are not always discrete and that there is considerable overlap in many of the projects as they bring together awareness-raising, legal change and plans of action and research.

While the specific strategies and outcomes vary, the underlying model of change of the sampled projects focuses primarily on awareness, advocacy and/or empowerment as means to the ultimate goal of reducing violence against women. This appears to be coherent with the UN's Human Security strategic concept of "protection and empowerment" that emphasizes a parallel focus on top-down protection and bottom-up empowerment measures.[8]

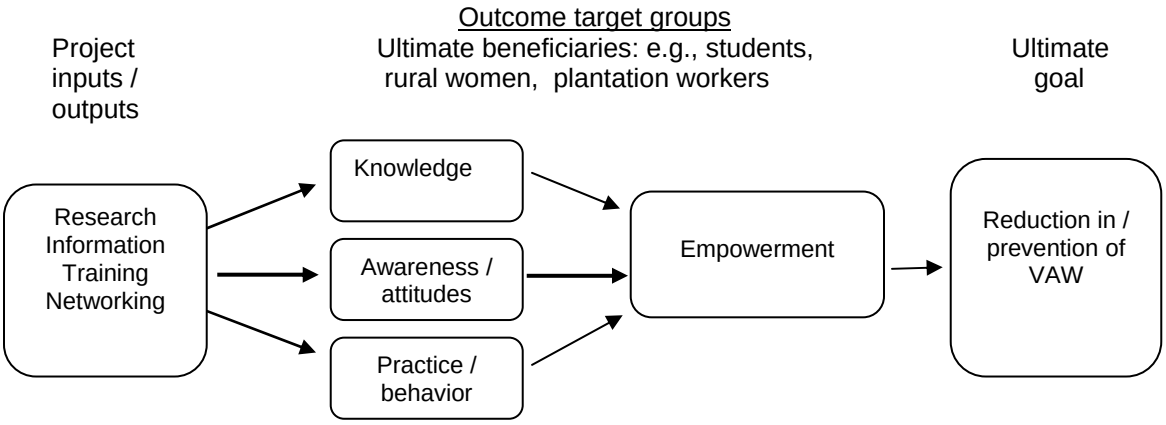
The direct intervention objectives of most of the projects focus on intermediate agencies or groups of actors within civil society through which change is sought at broader societal levels in order to reach the final goal of reducing violence against women. An alternative model is to work directly with the population within which ultimate change is sought. To facilitate conceptual clarity on the relationship between outputs and impact, the two models are depicted by the diagrams below. The two models work with similar inputs and outputs to reach the ultimate goals, but target intermediary agencies as vehicles to effect institutional change in model one, and the empowerment of the ultimate beneficiaries in model two. Model One shows the new project focus (for 2005 and beyond) on the implementation of laws and policies.[9]

Change model 1: Advocacy



[7] Advocacy refers to a range of activities including lobbying, influencing, campaigning, etc. Related outcomes refer to specific change objectives such as changes in policy, laws, implementation procedures, intervention protocols.
[8] UNTFHS Guidelines, ochaonline, 2006.
[9] As stated in a more recent revision of the Trust Fund strategy that applies to projects funded as of 2005 (UNIFEM, 2004).

Change model 2: Empowerment



In some cases the two models apply in a parallel fashion, such as when both students and high school officials are targeted for behavioral change.

The new generation of TF projects (2005 and onward) will focus specifically on countries that have already instituted laws and policies, promoting the effective implementation of these policies and monitoring the extent to which they have the desired effects. Against this evolution of project focus, the models presented above can be viewed as first generation models or necessary prerequisites to subsequent phases in a global UNIFEM strategy.

The evaluation of selected projects has been clearly mandated by the Fund's management, and the operational guidelines that have been distributed with the annual call for proposals dedicate a section to the need for impact assessment. The typical evaluation response comes in the form of a qualitative narrative, descriptions of activities and outputs, and/or simply the research document produced with the funds where research has been funded. None of the projects in the current sample provide evidence of a systematic evaluation.

The intent of the methodological guidelines is to facilitate rigor and consistency in the evaluation of project impact specifically or outcomes in general.

Purpose and structure of the guidelines

As the change models imply, Trust Fund projects share certain strategic generalities, but they also evolve in response to successes achieved and new obstacles encountered. Thus a new generation of projects will focus on the implementation of policies and laws in order to narrow the gap between protective mechanisms and the actual benefits realized by the ultimate beneficiaries. [10] The guidelines need to be responsive to the broader strategic interests of Trust Fund management, while providing practical support to grantees for the assessment of project impact or outcomes. Given a strong emphasis on process and the relatively short duration of the projects, an assessment of impact does not apply evenly to all projects.

[10] UNIFEM's New Strategy for the Trust Fund is provided in Appendix X.

Rather than determining project "evaluability" after implementation, the Trust Fund should assume a more assertive role during the project selection phase and distinguish between three types of cases: those that will be earmarked for rigorous impact evaluation, those that will conduct less rigorous outcome evaluations, and those that are best suited for process-focused evaluations. Funding for evaluation should be allocated accordingly.

The selected sample of past projects provides content and examples for the guidelines, which in turn are oriented toward future projects that are identified for impact or outcome evaluation according to TF decisions with regard to strategic evaluation investments.

It is in this vein that the guidelines are organized for two audiences, addressing the need for strategic decisions with regard to evaluation investments on the part of fund management on the one hand, and providing appropriate methodological guidelines for impact or outcome evaluation to grantees on the other.

The guidelines are thus presented in two parts. Part One clarifies the meaning of impact evaluation for the Trust Fund projects and examines the various options for evaluation design. It does so against the backdrop of UNIFEM's global and evolving strategy to reduce violence against women, and in light of the constraints imposed by the nature of small "innovative and catalytic" initiatives (short term, low cost) that characterizes much of the Trust Fund portfolio to date.[11]

Part two is intended for Trust Fund grantees and for associated technical support persons, possibly provided by UNIFEM. It provides phase-by-phase guidelines for impact and outcome evaluation, from project design through the presentation and utilization of the data. The intent of part two is to provide the basic options for each phase as a basis for making methodological decisions and as a platform from which each option can be explored in greater detail according to interest or need. Links to available on-line resources for detailed methodological information are provided at the end of the guidelines.

[11] While the future funding structure is expected to lift some of the constraints linked to budget and time, it is assumed that small initiatives will continue to make up a portion of the portfolio, in line with the original mandate.

PART ONE: DECISION GUIDELINES FOR UNIFEM

1 Impact versus outcome evaluation

While the original intent of the guidelines was to focus specifically on impact evaluation, discussions in preparation of the first drafts of these guidelines have indicated the importance of situating impact evaluation within a broader range of outcome evaluation options.

Program (or project) evaluation is "the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future programming" (Patton, 1997). Within this group of evaluative activities, impact evaluation can be further differentiated along a continuum of specificity and rigor. But as Bamberger points out, the meaning of impact evaluation is far from uniform to the various actors in international development (Bamberger, 2006).

As currently explained by the Trust Fund management to grant applicants, impact is an assessment that "evaluates to what extent (the) expected results were achieved and what these results mean to those involved in and/or affected by the project".[12] The UNDP guidelines for evaluation are also quite general, defining impact as the broad and longer-term outcomes of an intervention (UNDP, 2002). For other agencies yet, the term includes emphasis on certain qualifiers such as durability, sustainability and equitability.[13] These definitions leave the issue of methodological rigor fairly open-ended.

For the World Bank evaluation community, the objective of an impact evaluation is to determine a) whether a project has had the desired effects on individuals, households, and institutions and b) whether those effects are attributable to project intervention (Baker 2000; Wassenich & Whiteside, 2004; Ravallion, 2005).[14] Within the larger family of evaluations, impact evaluation is thus a special case charged with delivery of proof in order to become a defensible basis for modified policies, program extensions and replication. At its most rigorous, it follows the dictates of experimental design in social science research (the proverbial "gold standard") and becomes considerably demanding and costly.[15] To date, a relatively small proportion of World Bank projects are evaluated in accordance with such standards.[16]

Confronted with this type of reality, Rossi advocates the use of a "good enough" rule when designing impact evaluation: the evaluator should choose the best possible design from a methodological standpoint after having taken into account the potential importance of the results, the practicality and feasibility of each design, and the probability that the design chosen will

[12] Operational Guidelines for the Trust Fund in Support of Actions to Eliminate Violence Against Women, 2001.
[13] For example, CARE International's definition is "equitable and durable improvements in human wellbeing and social justice" (CARE 2003).
[14] Also referred to as "counterfactual analysis" by some (Ravallion, 2005).
[15] The recently estimated cost of rigorous impact evaluations for World Bank projects ranges from $200,000 to $900,000 each (OED and Impact Evaluation – A Discussion Note, 2005).
[16] The MIT Poverty Action Lab notes in 2004 that only 2% of all World Bank evaluations conform to the rigors of randomized control design (reported in the New York Times, July 2004).

produce useful and credible results (Rossi, 2002). Patton, on the other hand, emphasizes utility to the intended users (Patton, 1997).

Based on discussions with a number of peer reviewers and advisors that have reviewed the first draft of the guidelines, we propose the broader term "outcome evaluation", which includes impact evaluation as the technically most rigorous assessment of "net effects", and less rigorous variations where proof of causality is relaxed, but which nevertheless require a systematic approach to inquiry and data collection. Such a compromise is consistent with the conclusions of Valadez and Bamberger in examining evaluation options for projects characterized by constraints or for so called RealWorld evaluations (Valadez & Bamberger, 1994; Bamberger et al, 2004).

The guidelines will thus differentiate between the more rigorous classical form of impact evaluation that refers to the measurement of net effects and proof of causality, and outcome or summative evaluation that refers more generally to the end results of an intervention (as also described by Weiss, 98; Chen, 1990) and which relaxes the proof of causality requirement.

The decision to conduct a rigorous impact evaluation should be made before project implementation, and because of considerable requirements in technical and other resources, we would recommend that only a carefully selected percentage of the funded projects be earmarked for this type of evaluation. The evaluation of effect-level outcomes, on the other hand, should be an integral part of all future projects. The current guidelines will not lose sight of the possibilities for rigorous impact evaluation, but the focus of recommended practice will be on outcome evaluation in general, mindful of the constraints in time, funds and skills that characterize so many TF projects.

The following section discusses the different evaluation models that fall under the heading of impact and outcome evaluation. This, together with a review of the constraints encountered in Trust Fund projects, becomes a basis for developing a decision guide for matching project to evaluation design and for a reasoned allocation of evaluation resources.

2 Design options for impact and outcome evaluations

The purpose of impact evaluations is to find out whether a given intervention produces the intended effects or impact. According to Rossi et al, "the design of impact evaluations needs to take into account two competing pressures: on the one hand, evaluations should be undertaken with sufficient rigor that relatively firm conclusions can be reached; on the other hand, practical considerations of time, money, cooperation, and protection of participants limit the design options and methodological procedures that can be employed." (Rossi, 2002). In some instances, the available options are so limited that an assessment may not be appropriate.

Evaluation design determines the extent to which it can be shown that the project intervention has had the desired impact. Along a spectrum of rigor, the World Bank describes four categories of impact evaluation, of which two are considered sufficiently rigorous to deliver reliable estimates of impact attributable to the intervention. According to the distinction we made in Section 1 between classical impact evaluation and more general outcome or summative evaluation, the first two can be said to represent examples of the former type and the second two of the latter type of evaluation. That is, the last two categories can provide evidence of desired outcomes, but cannot reliably attribute these to project intervention.

1 Randomized design
2 Quasi-experimental design / matched control group design
3 Ex-post comparison of project beneficiaries with control group
4 Rapid assessment or review, conducted ex post

Randomized or matched control group design

Designs 1 and 2 fall into the category of rigorous impact evaluation. In the first instance, cases are assigned randomly to a study or intervention group and a control or non-intervention group as, for example, in clinical trials. This means that the evaluator collaborate with the project implementers from the outset, assigning cases as appropriate.[17] In the second case, comparison groups are created through reasoned matching, before, during or after project implementation. The technically most feasible matching option for the TF projects might be pipeline matching, where the baseline characteristics of a new project cohort serve as comparison reference for the cohort of a completed project.

Model 1 is promoted as the "gold standard" in terms of its rigor by many authorities on evaluation.[18] The implementation of this model involves considerable technical and managerial challenges, as Rossi, a top authority in evaluation, warns (Rossi, 2002). This is even more so the case in the field of international development, as, perhaps, evidenced by the small number of World Bank applications.[19] On the other hand, while such designs ensure much greater control of biases, they need not be any more labor-intensive and costly than other systematic data collection activities and can be used by NGOs especially in cases where they phase in programs over time. In such cases, "randomization will often be the fairest way of determining the order of phase-in" (Kremer, 2001). Not least, they are cheaper than pursuing ineffective projects and policies.

Randomized and quasi-experimental evaluation designs are feasible if planned in advance, but should be well supported by the Funding agency. As Kremer notes for randomized evaluations specifically, "Given that accurate estimates of program effects are a public good, and in fact, largely an international public good, randomized evaluations should be financed internationally" (Kremer, 2001).

Ex-post comparison of project beneficiaries with control groups

Ex post designs refer to impact evaluation designs undertaken subsequent to the delivery of the program to the intervention group. They include secondary analyses combined with the use of a quasi-experimental approach where it was not possible to construct the comparison group before project delivery, such as when the decision to do an impact evaluation comes after the project has begun. The comparison group has to be constructed retrospectively with the help of secondary data, recall, interviews with key informants and participatory techniques. Ex post

[17] The political difficulty of this requirement is described in …….
[18] There has been considerable controversy on this subject, between proponents of this view (National Institute of Health, World Bank) and opponents (American Evaluation Association, National Endowment for the Arts). For our purposes, we shall not enter into the complexities of the debate.
[19] According to one estimate issued in September 2004, only 2% of the World Bank's projects had been evaluated in conformity with this level of rigor.

matched control evaluation can also be achieved with statistical controls (Rossi, 2002), provided the sample is large enough. However, the use of statistical controls or of multivariate statistical models requires a fairly high level of technical expertise, thus this design is not appropriate where access to technical skills is limited.

Rapid or other ex post outcome assessments

When time and cost are of paramount concerns, rapid and economical ways of collecting ex post data may be the only option available for outcome assessment.

Rapid outcome assessments use a variety of techniques that are relatively quick to administer, such as key informant interviews, group interviews or focus group discussions, secondary data, review of records, and direct observations. Unfortunately, the simplicity of development and administration is too often countered by the difficulty of interpretation and analysis of the data at the end. Qualitative data coming from open-ended questions must be subjected to a very laborious process of extracting and interpreting key themes and patterns if it is to make sense. The results are likely to be primarily descriptive and may not be very conclusive from an impact perspective.

Rapid assessments are generally better suited to the interests of project implementers. When combined with participatory methods, they facilitate ownership of evaluation findings among beneficiaries and staff and become a basis for community- or project-driven project revision. In such cases, the principal interest is more likely to be in process rather than impact or outcome evaluation.

Within the family of ex post outcome assessments, the process of outcome mapping developed by the International Development Research Centre (IDRC) provides a low cost methodology that is relatively easy to manage with local resources.

An alternative or complementary approach is the case study, conducted by an expert in the subject matter. However, this option does not correct for the inherent weakness of rapid or qualitative designs: the absence of objectively verifiable reference points, temporal and/or across comparison groups.

Another, more conventional, approach for delivering outcome results within this last category is ex post cross-sectional surveys. Such surveys can be rendered more rapid by eliminating all but the most essential subjects from an evaluation instrument or by reducing time-consuming questions such as those on income, expenditures, time use, etc. (Bamberger, 20004).[20] An ex post survey, complemented with qualitative techniques of inquiry, can answer important questions that are central to the outcome question, as well as provide contextual and explanatory detail from different perspectives, thus making a convincing case for the extent of project effectiveness.

3 Constraints to impact or outcome evaluation of Trust Fund projects

[20] Unfortunately, the focus of TF projects on inherently qualitative outcomes such as awareness, attitudes, knowledge and capacity may not lend itself readily to the pruning of survey instruments.

Referring to the implementation of rigorous impact evaluation in developing countries, including cases such as Community-Driven Development projects that have been added to the World Bank portfolio in recent years, constraints to the implementation of such evaluations have been described by a number of authors (Valadez & Bamberger, 94; Bamberger, 2004; Wassenich & Whiteside, 2004). The following list adapts their observations to projects in the Trust Fund portfolio.

Budget constraints: The average budget of the TF grants reviewed is $50,000. With 10% set aside for evaluation expenses, grantees work with an average evaluation budget of $5000 and in some cases less. In a context where technical support is rare and costly and communication and transportation infrastructures are limited, such amounts clearly exclude a range of evaluation options, in particular those involving a high level of technical expertise (e.g., for sampling and other methodological decisions and analysis).

Time constraints: The average duration of TF grants is one year, with only few exceeding that period (4 out of 15 in the selected sample). This often precludes the observation of desired change beyond the output level. On the other hand, the duration is consistent with the Funds emphasis on low-cost, innovative and catalytic projects. According to the criteria that guide project selection, the projects "should seek to put in place mechanisms that contribute to the continued involvement of target groups and/or partners beyond the life of the project." Moreover, the true impact of projects having advocacy-related objectives may take years to become reasonably accessible to measurement.

Capacity constraints: The selection of control or comparison groups, whether through use of randomization or a matching process, requires a professional level of expertise that is rarely available at the local level, whether at the time of project/evaluation design or at the time of the final evaluation.

Data constraints: In cases where the evaluator is called in toward the end of the project, access to appropriate data can be problematic. Secondary data can be difficult to find, especially in post conflict zones where data have been destroyed. Data from large national or regional surveys may not be available in sufficiently disaggregated form. Baseline data, when available at all (a relatively rare occurrence in the experience of this evaluator), may be inappropriate due to revisions in project strategy or to technical shortcomings in the study. Staff turnover can limit the availability of institutional memory.

Skill constraints: The review of the selected projects provides evidence of only modest levels of evaluation skills. None of the evaluation designs have the basic elements of good evaluation design maximizing validity and replicability. For example, It is rare to find clearly stated measurable impact indicators that would facilitate an impact assessment, and rarer yet to find baseline inquiries that would permit a before-and-after comparison of the target population. None of the selected projects proposed the use of control groups or areas. Future projects targeted for impact evaluation will need technical assistance not only by means of guidelines, and but in the form of an evaluation professional that can provide hands-on support.

Constraints in controlling the link between project implementation and evaluation: The Trust Fund responds to proposals from field-based organizations. The proposals are selected by a

panel on the basis of several criteria.[21] It is assumed that the selected proposals reflect appropriate contextual concerns due to the field experience of the applicants. The link between project design and evaluation design tends to be weak, and evaluators, when used, are usually called in at the end of the project when the extent to which the link can be strengthened is limited. The existence of unreported other social programs and intervening agencies can present another constraint in controlling conditions for impact evaluation.

Complexity: Complexity issues requiring a more advanced level of skills arise from several quarters. The main culprit is poor project design which can increase enormously the complexity of evaluation at the end of the project. Certain outcome measures are inherently difficult, especially those that are contextually sensitive as are awareness, knowledge and attitudes. As has been noted for CDD interventions, such projects "cannot escape the importance of context" (Wassenich & Whiteside, 2004). Another source of complexity is the difficulty of determining cause and effect in advocacy initiatives and outcomes, especially when such initiatives attempt to influence more general attitudes and values in society (Chapman, 2001).

While the guidelines are based on a sample of past projects (funded from 1997-2003), the intent is to provide guidance to new projects with the assumption that the evaluation will be integrated from the outset.

4 Review of key issues for the evaluation of Trust Fund projects

UNIFEM's criteria for the post-2004 generation of Trust Fund grants anticipate larger and longer grants, which promise to address an important set of constraints imposed by time and funding on the evaluation of past projects. The guide for selecting appropriate evaluation designs looks forward to a new generation of projects that eases the time and budget constraints for impact evaluation, and uses a sample of past projects for design alternatives.

Time and budget constraints

Time and budget constraints, clearly related, are key determinants in deciding upon evaluation design. The allocation of 10% of the project budget to evaluation is entirely reasonable and in line with the general recommendations within the evaluation and donor communities. The bottom line in the case of the projects reviewed is that the great majority consists of low budget initiatives that are responsive to local conditions and that may be highly effective, but that do not lend themselves well to demanding and costly evaluation designs. Our review of the sampled projects also shows that budget allocations are not clearly correlated with change objective or strategy type. Thus a first filter for determining evaluation design is the practical budget constraint, with the closely related constraint of time being a second one (depending on the ability to conduct follow up monitoring and evaluation).

[21] One of the criteria suggested is organizational strength which includes the experience to implement the project, with preference given to organizations that are based in the country. (see UNIFEM's suggested criteria for evaluating TF proposals, UNIFEM Trust Fund Management (Appendix 2).

Table 1: Projects by budget allocation, change objective and strategy type

Budget-based group (project #)*	#	Principal change objective	#	Key strategy type	#
< $5000 (4,6,7-10, 12)	7	Awareness/knowledge (A/K) A/K + advocacy A/K + advocacy + empowerment A/K + advocacy + service provision	1 3 2 1	a) Education / training / capacity building b) Research & documentation c) Networking / partnership building a + b a + c a + b + c	0 1 0 3 1 2
$5000-$6500 (1-3,5,11, 13,14)	7	Awareness/knowledge (A/K) A/K + advocacy A/K + advocacy + service provision A/K + capacity + empowerment	1 4 1 1	a) Education / training / capacity building b) Research & documentation a + b a + b + lobbying	4 0 2 1
$10000 (15)	1	A/K + advocacy + service provision	1	a + b	1
Total # of projects in sample	15		15		15

* See corresponding number in the list of projects in Appendix 1.

Aside from time and budget considerations there are other factors need to be considered when singling out projects for the more rigorous impact evaluation versus other types of evaluations. We group these under three headings: general, methodological and practical considerations. They apply to any type of evaluation design.

General considerations

- **Project design**: A basic rule of thumb that is proposed to help determine whether a given project justifies an impact evaluation at all is that the program's objectives must be sufficiently well articulated to make it possible to specify credible measures of the expected outcomes (Rossi, 2002). This condition is predicated upon a strong and clearly articulated project design. The project proposals reviewed suggest that most of the grantees will need considerable support to provide the clarity needed for good evaluation design.

Future project proposals will be submitted as project outlines rather than detailed project plans (UNIFEM, 2004). The decision for impact evaluation may thus need to be made on the basis of other considerations. Once selected, the grantee should agree to a systematic support agenda to review the project design and ensure that it provide an appropriate basis for evaluation design in general and for impact evaluation in particular.

- **Possibilities for post-project evaluations**: One valuable feature of Trust Fund projects is a social network context that facilitates durable relations and the possibility of follow up inquiries after the end of project funding. Where evaluation conducted within the fund cycle maybe constrained by time, the broader relational context may permit retrospective reviews and post project follow up impact monitoring (in the case, for example, of advocacy projects and the publication of new research).

- **Stakeholder interests and grantee buy-in**: Key interests in evaluation should be clarified and sorted out during the proposal selection process. The demand for rigorous impact evaluation comes rarely from the grantee side. It is therefore up to UNIFEM to take the lead on

what type of evaluation should be undertaken, determine how much support and funds should be allocated, and lobby for buy-in from the grantee.

- **Technical support and capacity building**: Evaluation initiatives should provide capacity building at the local level through a learn-by-doing process, so that grantees will derive tangible value in the form of learning and documentation. A mixed methods approach (discussed in Section 3) will further enhance opportunities for learning at many levels. Where grantee cooperation can be secured for a rigorous or systematic impact evaluation, technical support must be provided from the outset, with well planned follow up activities that facilitate both project monitoring and the conditions for impact evaluation, and stable communication procedures that ensure continuity and effectiveness of the collaboration.[22]

Methodological considerations

- The key change objectives of Trust Fund projects are awareness, usually combined with some evidence of follow up action and behavioral change with the possibility of continuity. These are particularly challenging areas to evaluate with conventional techniques (Chapman (2001) Because of the importance and complexity of contextual characteristics, the use of a mixed methods approach will be essential, not only for evaluation, but also for project monitoring. Agreement must be obtained that the grantee will integrate such monitoring procedures into ongoing project activities and record the data on a routine basis.

- Awareness and advocacy work usually operates at multiple levels. The project proposal must indicate clearly who the target population is at different levels, and what the possibilities are at each level. Options for selecting control groups should be considered when target groups are identified. A project where target groups cannot be clearly delineated, and/or where control groups cannot be envisioned, is not a good candidate for impact evaluation.

Reasons for the difficulty of identifying target groups may be due to the project's exclusive focus on output with no project control over where and how the output related information is received. This is the case sometimes with a piece of research or publication where funding ends with the production of the document.

- The outcomes of advocacy work frequently lead to compromise rather than outright victory, and objectives may shift as perceptions adapt to emergent realities (Chapman, 2001). It is important to define what constitutes satisfactory results within a range of possible outcomes as a basis for the development of appropriate success indicators. The setting of clear and realistic objectives is important, while setting the sights as closely as possible on desired outcomes.

- The role of cooperation and networking in awareness/advocacy projects can render the use

[22] The consultant's experience shows that crisis prone local conditions can block or interrupt communication, but sometimes do so more readily than is necessary.

of conventional impact evaluation techniques more complex, especially with regard to the attribution of effects. The selection of study and control groups will need to be based on an analysis of social networks involved.[23]

- While most of the project proposals distinguish between short-term and long-term goals, the possibility of monitoring medium and long term impact beyond the funded period should also be considered.

- Awareness and advocacy work can involve political and individually damaging side effects that must be anticipated with the help of tools such as benefit-harms analysis (see section xx in Part II).

Questions to determine candidacy for impact or outcome evaluation

Taking these methodological considerations into account, the following questions should be asked to determine whether to invest in rigorous impact evaluation versus other evaluation designs.

1 Are project objectives clearly stated, and if not, can they be revised according to SMART criteria (see Section 2 in Part II)?
2 Are project objectives demonstrably aligned with Trust Fund strategy and goals for change?
3 Are project objectives realistically achievable within the funded project period? If not, can they be evaluated after the project ends?
4 Are the key target populations clearly identified and delineated such that they provide a realistic and technically sound basis for a study sample and/or a control sample? What are the possible sampling frames that can be used (see Section 4 in Part II)?
5 To what extent is the grantee team committed to cooperating with recommended evaluation protocols?
6 How will the evaluation results be used by a) grantee; b) UNIFEM? What policy questions can they help to answer? What will be their visibility and reach?
7 Can the project be replicated with or without funding from the Trust Fund? What, if any, are the follow-up phases considered?
8 What technical support resources are available locally, regionally, globally? What is the most appropriate technical support strategy?
9 What are the expected costs of a rigorous impact evaluation?

The following sub-section attempts to provide some guidelines for answering questions 8 and 9.

Practical considerations for the implementation of evaluations

Technical support strategy

[23] One recently developed subcategory of evaluation approaches focuses on network analysis, a primarily descriptive and inductive tool that can complement other evaluation techniques in complex social environments (Davies, 2003).

Given the cost and resource requirements of rigorous impact evaluation, only a small and select number of projects should be earmarked, especially during the first support year. A second group would qualify for less rigorous outcome evaluation, using a mix of conventional quantitative and qualitative techniques. For a third group, a relatively modest mix of monitoring procedures and qualitative exit interviews might be most appropriate.

Once projects have been sorted according to an evaluation priority scale, a technical support strategy needs to be adapted to the recommended level of expertise required on the one hand, and to the availability of technical capacity at local or regional levels on the other.

a) For impact evaluation:
Impact evaluation specialist with sampling expertise (or else access to consultation with sampling specialist), experience in mixed methods approaches, and specialized IT support (familiarity with data entry and statistical software).

b) For outcome evaluation:
Evaluation specialist using Trust Fund evaluation guidelines and web-based resources, supported by IT person for data entry and processing.

c) Monitoring and qualitative exit interviews:
Local consultant working jointly with project-based management and staff.

The technical support strategy should ensure that impact and outcome evaluation events are exploited as fully as possible as opportunities for capacity building through learning-by-doing.

The long term strategy of technical support is to a) develop and test a set of basic models for rigorous impact evaluation for selected Trust Fund projects; b) build capacity to supervise outcome evaluations at the regional Trust Fund level; c) build local evaluation capacity.

Using a team approach based on a mix of carefully selected local and regional resources under the guidance of an evaluation expert is the most effective way to build capacity and keep costs to a minimum. Opportunities for capacity enhancement are always desirable at local levels, and institutions of higher learning are usually quite willing to collaborate in such endeavors at reasonable fees.

Models of technical support that have worked particularly well in my experience are those that begin with a small core group of evaluation agents led by one evaluation specialist that proceed to decentralize capacity through hands-on training in the field. Baseline evaluations usually provide an excellent opportunity for capacity building at the local level. In two cases, the original evaluation expert was phased out after a period, while the team continued on its own, with the possibility of distance-based support and/or supervision.

Approximate cost considerations

Impact evaluations are notoriously expensive, though the cost depends on the scope and complexity of the evaluation. We agree with Kremer (2003) who claims that when well planned, the cost of impact evaluation need not exceed that of other systematic outcome evaluations.

Human resources

The key cost item is the expertise that ensures a defensible evaluation result. If an evaluation expert is recruited from the open market on a contractual basis, the fees should be based on the going rate as set by USAID standards, adjusted for experience and the consultant's past rates.

An impact evaluation (designs 1-3 in Section 2 above) will typically require 30-50 consultancy days, from evaluation design and preparations to the delivery of the evaluation report. As a rule of thumb, the consultant's expenses will constitute slightly more than half of the total cost of the evaluation (without travel cost to and from the site). There is more flexibility in

Other human resource costs include: a logistics coordinator, an IT supervisor, secretarial support, data collection team, 2-4 data entry agents, drivers, and other staff at critical points (planning stages, reviews, etc.).

Training & staff development

Training is one of the most critical activities, especially when working with local recruits. The costs involve workshop facilities and logistics, materials, and transportation for field-based exercises.

Data collection, processing, analysis & reporting

Includes hire/use of equipment and facilities, supplies for the printing and distribution of data collection instruments and reports, supplies for the field work, occasional purchase of software.

Travel & accommodation

These costs vary considerably by location, but can be substantial and should be carefully planned in advance depending on the terrain, distances, access to fuel, lodging facilities.

Overall cost estimates

Based on past experience in various locations in Africa where costs tend to be more elevated, the estimated cost of an evaluation that includes a household sample of 500-800 plus qualitative information ranges from $35,000 – $50,000. Smaller samples will reduce cost.

For rapid, qualitative surveys, costs can be as low as $1000.

Conclusion

The Trust Fund projects are representative of an increasing prominence of rights based approaches in international development. While instruments and methods exist for the measuring of impact of service delivery oriented program, the evolution of equivalent tools for measuring awareness and advocacy related initiatives remains in its nascent phase (Kumar, 2005). In view of the fact that such initiatives often involve broad mobilization and reach before they can achieve observable results, the short-term and low-cost nature of the TF portfolio makes the task more daunting yet.

No matter how challenging, the effort is launched and if the guidelines are not able to provide precise procedural directions, they lay the groundwork for the reflections that are prerequisite to well-reasoned evaluation decisions in the complex, adaptive environments of human rights initiatives.

PART TWO: METHODOLOGICAL GUIDELINES FOR IMPACT EVALUATION

1 Introduction

The second part of the guidelines is written for the grantees and assumes that some technical support will be provided by UNIFEM when projects are earmarked specifically for impact or outcome evaluation.[24]

UNIFEM supports two types of evaluations for Trust Fund grants: 1) Impact evaluation; 2) Outcome evaluation. The objective of an impact evaluation is to determine whether a project has had the desired effects on recipients, and whether those effects are attributable to project intervention. An outcome evaluation refers more generally to the end results of an intervention and relaxes the proof of causality requirement.

Definition box 1

Impact evaluation:
The focus of impact evaluation is on the effects of the project on recipients. The intent is to find out whether an intervention actually produces the intended effect.

Whether impact or outcome evaluation, information required for the evaluation includes:

1 Realistic outcome objectives, that is, objectives that are concretely attainable within the project's time frame with available resources.
For example: *By (month, year), at least xx percent of all cases involving domestic violence will be processed according to the (new law).*
And NOT: *By the end of the project, the (new law) will be applied to all cases of domestic violence.*

2 Clearly delineated target groups of beneficiaries or recipients.
For example: *The target beneficiaries are women community leaders in ten regional centers*
And NOT: *The target beneficiaries are women in rural communities in Ethiopia*

3 Objectively verifiable indicators.
For example: *Number of mechanisms designed and applied to provide for the protection of women.*
And NOT: *Increase in women's protection*

Good baseline data is highly desirable, but the extra cost is not always recommended in projects that are adaptive to emergent conditions as projects with advocacy objectives often are (Chapman, 2001).

[24] It is important to distinguish between impact evaluation which focuses on the effects of a given intervention and process evaluation which answers questions about program operations and implementation. While the guidelines indicate where process evaluation might be the most feasible course, their methodological focus will be on impact or outcome and not on process evaluation.

The design of any evaluation should begin with the design of the project itself, for "once a project has begun, many options for making the evaluation more rigorous are no longer available" (Wassenich & Whiteside, 2004).

The first task of the evaluator should be to review the project design in order to obtain the relevant information, either retrospectively or as a first step in the development of an evaluation plan.

2 Project design

The basic elements of project design include the project hypothesis, a set of logically coherent objectives and a strategy that links inputs to the achievement of these objectives, given a clearly stated set of assumptions about existing possibilities and constraints. Project design in turn provides the basis for the identification of objectively verifiable indicators, the key ingredient for project evaluation.

Definition Box 2

Project Design:
The planning of desired outcomes on the basis of the identification and prioritization of problems and coherent linkages between inputs, outputs and results

The first step for project design is a clearly articulated project hypothesis that shifts the focus from the analysis of problems to the potential for solutions.

Definition Box 3

Project hypothesis:
The presumed correlations between cause and effect, and between intervention and impact that describe a path for change.

The construction of a project hypothesis should be a participatory process and can be achieved through the following basic steps.

- State the problem to be addressed and its causes and effects. A problem tree can be a very useful participatory tool for this phase. A description of the method is readily accessible via the internet (e.g., via www.info.worldbank.org).

  Note that causes should be organized to differentiate between fundamental structural causes embedded in social and cultural systems that may be more difficult to address but that will influence outcomes, and behavioral and attitudinal causes that are typically the target of the Trust Fund projects.

- Convert the problem and its key causes to solutions or objectives and anticipated outcomes.

- Develop a visual diagram linking the main problem to its possible solutions through planned interventions.

- Identify assumptions concerning possibilities and constraints.

Project Objectives

> **Definition Box 4**
>
> **Project Objective:**
> A project objective is a specific description of an intended outcome (NSF Handbook, 2002).

Project objectives should be stated in accordance with a project hypothesis and an implementation strategy that shows how planned outputs produce effects and impacts at broader and sustainable levels. An excellent tool for facilitating project design is the logic model, also known as the logical framework approach (LFA). [25] Based on cause-effect logic, such tools offer a systematic way to examine and present linkages between resources, activities, and the desired changes or results. The objectives statement should include what is to be achieved and when and should clearly identify the direct target population – both, at the level of intermediary agency and at the level of the ultimate goal (linking project objectives to the Fund's global objectives).

Outcome objectives should be stated for the different levels of an implementation strategy from immediate result to long-term goal. They should be described as desired states to be achieved, rather than as processes, to avoid the confusion between objective and activity that is so frequently found in Trust Fund proposals. They should distinguish clearly between immediate outputs, expected or desired behavioral changes, and longer term outcomes, a distinction that is especially critical for the design of an impact evaluation.

No matter at what level they are stated, outcome objectives should be visualized and stated as concretely as possible and include the following key elements:

- The expected observable change or effect (what will be different?)
- The identification and number of intended beneficiaries (among whom will the difference be observed?)
- The time period within which change is expected (at what point in time will the difference be observed?)

To encourage the visualization of a future desired state of being rather than a process, I recommend that the objectives statement begin with the point in time by which a given objectives is to be reached. For example, "by (date), (concrete description of outcome)".

Outcome objectives that provide a sound basis for evaluation design should be SMART:[26]

Specific: is the objective clear about what, how, when and where the change is to occur?
Measurable: what is the evidence of change and of the achievement of objectives?
Achievable: is it reasonable to expect the desired change within the designated period given resources and constraints?
Realistic: do the objectives address the scope of the problem and propose reasonable programmatic steps?
Time-related: within what time period is the objective to be attained?

[25] Guidance to the use of LFA tools can be found through the following websites: www.wkkf.org/pubs/tools/evaluation; www.uwex.ed/ces/pdande/evauation; www.jiscinfone.ac.uk; www.ifad.org/evaluation/guide; www.web.mit.edu.
[26] The SMART method for evaluating objectives was introduced by Peter Drucker and the Management by Objectives movement of the fifties.

An additional quality to be considered for an outcome objective is its relevance to donor-based goals. It should be clear how the project-specific objective aligns with the broader goals of the funding agency. Thus an outcome objective should flow from a set of activities and outputs on the one hand, and contribute to the achievement of donor-specified goals in the broader scheme of things on the other.[27]

It is through this type of exercise that a more realistic assessment of achievable outcomes can be obtained, provided it draws upon the input and experiences from actors at different levels of the project implementation process. Thus it is important that the process be participatory.

Table 2.1 provides an example of how the results of an objectives tree can be transformed to produce objectives conforming to SMART criteria. The example provides the different levels of a hierarchy of objectives, which in turn presents the broader strategic context within which specific project objectives that are reasonably attainable can be situated.

Table 2.1: Example of SMART objectives

From objectives tree	*SMART objectives for project design*
A. Immediate objectives / results	
Increased knowledge of legal rights and strategies among rural women	By (month/year), at least (number) of rural women clients of (x) regional training centers will be able to demonstrate knowledge of their legal rights and on possible strategies for survival and protection
Increased knowledge of women's rights among decision makers in education, government and labor unions.	By (month/year), at least (number) of teachers in (x) number of schools, union officials and ministry officials within (x) regions will be able to demonstrate knowledge of women's rights
B. Intermediate effect objectives	
Increased use of FIDA-based legal services by women.	By (month/year), women clients of FIDA-based legal services in the regional centers will have increased by at least 15%.
Increased number of institutional collaborations to advocate legal reform.	By (month/year), interagency collaborations with objectives to defend women's rights across the 10 regions will have increased by at least 20%.
C. Longer term impact objectives	
Increased policies and advocacy initiatives favoring women's rights in educational institutions and plantations.	By (month/year), initiatives in support of protecting women from sexual violence and harassment will have been instituted in at least 30% of the schools and in 50% of the plantations.
D. Ultimate goal	

[27] One practice is to state the broader global goal at donor level as part of the context within which project-specific objectives are articulated. It does not mean that the evaluation should attempt to measure the attainment of broad visionary goals, but that it keep the larger context within conceptual purview for stakeholder appreciation.

Reduced incidence of gender based violence in educational institutions / plantations.	By (month/year), the incidence of GBV in educational institutions and plantations will have been reduced by at least 30%

The list of revised objectives becomes the basis for the development of outcome indicators discussed below. But first, a very critical element within the formulation of objectives is the clear identification of target groups.

Definition Box 5

Target population:
A target population is a set of individuals or institutions expected to benefit from the proposed intervention

Target population

The identification of the target population should be as precise as possible, specifying who is included and who is not in terms of geographical radius and/or or any other critical characteristic (age group, sex, etc.). This identification permits to develop an appropriate sampling strategy.

One practical challenge for the design of impact evaluation is to identify the primary target group when the intervention operates through intermediary agents and cannot reasonably vouch for observable effects at the level of the ultimate beneficiaries within the funded time frame. For example, an advocacy initiative might envision improvements in the prevention of VAW, but for practical purposes, its target group consists of decision makers within intermediary agencies such as the legal system. Moreover, in this last example, the target group may change as the initiative is adapted to a changing political and/or institutional environment.

For projects that aim at outcomes at the attitudinal or behavioral level, as many TF projects do, the question should focus on the specific locus of change in knowledge and practice.

The questions that can guide TF project proposals when identifying the target population might be as follows:

a) When raising awareness is the proposed change objective:
Who are the most direct recipients of the awareness initiative?
What are their specific characteristics?
Whose awareness is the most critical in advancing the overall TF objectives? What are their characteristics? Where are they?
What attitudinal or behavioral change is sought through raising awareness?
Among whom?

b) When knowledge through new research and documentation is the proposed change objective (for example in the case of projects # 3 & 4, Appendix 1):
Who are the most direct recipients of the information initiative?
Whose knowledge is the most critical in advancing the overall TF objectives?
What attitudinal or behavioral change is sought through increased knowledge, and among whom?
What action is sought as a result of increased knowledge, and among whom?

c) When increased capacity is the proposed change objective:
Who are the most direct recipients of the capacity building initiative?
Who are the ultimate beneficiaries of the capacity building initiative?

Whose capacity is the most critical in advancing the overall TF objectives?
What action or change is sought through increased capacity, and among whom?

d) When increased advocacy is the proposed change objective:
Who are the intermediary agents whose advocacy skills will be improved?
Who are the decision makers targeted by the initiative?
Who are the ultimate beneficiaries of this initiative?

If an experimental model of impact evaluation is to be considered, it is here that the process of considering possible comparison groups, whether selected randomly or non-randomly, should begin.[28] The process of selecting individuals, groups, institutions or areas that somehow lie outside the intervention sphere can be politically sensitive and needs to be carefully discussed with the implementing agency. The subject of comparison groups will be further discussed in Section 4 below.

Outcome indicators

Evaluation indicators are signals or symptoms of a phenomenon of interest that are used to verify whether a proposed change has occurred. They can be (often are) loosely stated as concepts. The art of developing useful indicators for evaluation is the bridging of the gap between loosely stated concept and objectively verifiable measure – not an easy task as many know.

As with project design, the identification of indicators should be a participatory process, and is usually best when facilitated by an external program design or evaluation support person.

Definition Box 6

Evaluation indicators:
are signals or symptoms of a phenomenon of interest that are used to verify whether a proposed change has occurred.

Indicators of change should distinguish between outputs and outcomes. Where output indicators are typically used for the monitoring of project implementation (number of training sessions held, persons trained, etc.), outcome indicators (effect and longer term impact) are fundamental to the evaluation of impact.

Within outcomes, we distinguish between immediate and intermediate effects (change in behavior or practice) that are required to reach an ultimate impact goal, and the longer term impact. In the case of advocacy-focused projects, we might distinguish between effects at the level of organizational behavior or activism and impact at the level of policy change or policy implementation. At the ultimate impact level, we would expect evidence of impact on people's lives. TF projects, being relatively short-term, focus typically on change at the effects level.

We propose two phases for the development of indicators: the conceptual phase and the measurement phase. The process of identifying change objectives becomes the basis for proposing indicators formulated at the broader conceptual level. A sample of such indicators,

[28] The experimental model is described in the Decision Guidelines for UNIFEM and would require additional technical support.

applicable to TF projects having knowledge, skills and advocacy objectives, are provided in Table 2.

Table 2: Example of indicators at effects level at longer term impact level

Indicators of intermediate effects	*Indicators of longer term impact*
Applications of legal skills by trained women and community leaders Change in written publications Increased public discussion of issues related to women's rights Changed attitudes, knowledge, practice among: law enforcement agents social workers medical workers judicial staff teachers community leaders Increased number of advocacy-related initiatives Utilization of research results by national decision/policy makers Increased knowledge of women concerning their rights Increased number of discussions on gender equality in schools Increased evidence of networking among women's organizations University courses on gender equality Number of requests for further gender-focused training	Increased number of cases brought before court Changed policy / legislation (making reference to study findings) Changed policy implementation Decreased violations of women's rights by: law enforcement agents social workers medical workers judicial staff teachers community leaders Increased number of programs aimed at meeting the needs of women Increased number of programs addressing domestic violence and inequality issues Changes in family law favoring women's status Number of women participating in self-governance bodies at local level National structures dealing with domestic violence established National monitoring systems established to register cases of violence against women

For the second phase, the proposed concepts are redefined as measurable elements of evidence.

Good impact evaluation indicators should be:

- Measurable and objectively verifiable.
- Technically feasible, given available skills and resources
- Reliable, producing the same results across different measurements (by different observers, for example).
- Valid for the phenomenon in question, that is, measuring what they intend to measure.

- Relevant to project objectives at the appropriate level in the hierarchy (e.g., output vs. outcome).
- Sensitive to changes in the situation being observed.
- Cost effective, justifying the time and resources needed for the collection of the data.
- Timely, enabling the collection and analysis of the data within a reasonable period.

Using objectives from Table 3 and applying the criteria above, the associated indicators would look something like this:

Table 3: Example of objectives and indicators

Objectives for project design	*Indicators*
A. Immediate objectives / outputs	**Results or output indicators**
By (mm/yy), x number of rural women in x regional centers will have been provided with information and advice on their legal rights and on possible strategies for survival and protection.	Number of women having received information and/or advice on their legal rights and on possible strategies for survival and protection
By (mm/yy), x number of teachers, union officials and ministry officials will have received information and training on women's rights and gender issues.	Number of teachers, union officials and ministry officials having received information and training on women's rights and gender issues.
B. Intermediate effect objectives	**Effects indicator**
By (mm/yy), women clients of FIDA-based legal services in the regional centers will have increased by at least 15%.	Percentage increase in the number of women clients of FIDA-based legal services between Oct.97 and Oct.98.
By (mm/yy), interagency collaborations with objectives to advocate for legal reform with regard to women's rights in each of the 10 regions with FIDA interventions will have increased by at least 20%.	Annual percentage increase in interagency collaborations advocating for legal reform with regard to women's rights in each of ten regional centers since Oct.97.
C. Longer-term impact objective	**Longer term impact indicator**
By (mm/yy), initiatives in support of protecting women from sexual violence and harassment will have been instituted in at least 30% of the schools and in 50% of the plantations.	Annual percentage increase in the number of schools & plantations with initiatives in support of protecting women from sexual violence and harassment since Oct.97.
D. Ultimate goal	**Ultimate impact indicator**
By (mm/yy), the incidence of GBV in educational institutions and plantations will have been reduced by at least 30%	Percentage decrease in the incidence of GBV in educational institutions and plantations (controlling for initial increase in reporting).

The development of good indicators is laborious and there is no shortcut or easy recipe for arriving at the right indicators. It is essential that the endeavor be a group process and that no indicator be accepted without thorough group discussion. And it is a fine moment when after much discussion an indicator finally sits well with all concerned.

While indicators of expected outcomes are core to the evaluation process, a word of caution is in order about the sole reliance on predetermined indicators. By definition, they focus on expected change and agreed-upon areas of change, leaving aside the unexpected and unintended (Roche, 1999). A review of assumptions and of possible threats to the change hypothesis can stimulate thoughts outside the box and help identify contextual pieces of information that could throw light on areas outside the focal point of proposed outcomes.

Assumptions

Definition Box 7

Assumptions:
are the circumstances or conditions important for the success of the projects but beyond direct control by the project.

Assumptions should be carefully examined to determine:

- the importance of external factors that lie beyond the control of the project to the success of the project;
- the possibility of bringing certain external factors under the control of project management;
- the availability of sufficient information on external factors for monitoring and/or statistical controls;
- the analysis of counterfactuals for the evaluation of net effects.

Assumptions can be identified through the following process: [29]

- Identify external factors recognized as causes during problem analysis but not addressed directly by the project.
- Identify external factors not recognized in the project hypothesis logic but important to the success of the project.
- Assess the importance of each remaining external factor and the likelihood of it being realized.
- Determine if certain external factors can be brought under the control of project management and revise project design accordingly.
- Determine if information on external factors is both critical and obtainable.[30]
- Evaluate the threat of each external factor to the achievement of project objectives.
- Adjust the project hypothesis accordingly.

[29] Adapted from Caldwell, 2002: Identifying Assumptions and Key Questions. p.76.
[30] This information will be useful for counterfactual analysis, i.e., the development of a hypothetical scenario which supposes the course of action and outcomes without adoption of the intervention.

Table 4: Example of identification of assumptions

Question/step	*Description / explanation of factor*	*Importance*
1 External factors recognized as causes during problem analysis but not addressed directly by the project	a) Culturally & socially institutionalized male dominance in households and the public sphere b) Socially institutionalized economic poverty of women	a) A contextual given that lies beyond the immediate scope and capacity of this project, but that influences its implementation and may mitigate its outcomes. b) Economic exclusion and the resulting poverty of women threaten their capacity to act on their own behalf. While this can affect their ability to use FIDA-based services, the project itself can improve this condition. The project should ensure that the effects of this factor are minimized by facilitating the use of services for the poorest women in particular (poor in resources and in available time).
2 External factors not recognized in the project hypothesis logic but important to the success of the project	a) Plantations remain operational at about present capacity. b) No major strikes in educational system and plantations c) Political stability at regional level. d) Good cooperation from government and other officials. e) Male resistance in certain sub groups (religious or ethnic)	a) If plantations reduce operations, union officials will be concerned with overall survival rather than gender issues. b) During major strikes, gender-related issues tend to take a back-seat in priority. c) Political instability can limit access to and initiatives at the regional centers. Overall: gender-related initiatives require a relatively stable environment with no major threats to the basic survival of men. d) Good cooperation with government and other officials is essential to this project and must be actively promoted as part of project activities. e) Women in certain sub groups may experience heightened threats within their households, families and communities due to their participation in the project. A benefits-harms analysis should be conducted (see next section below).
3 Other external factors that may influence the outcomes of the intervention	a) Competing initiatives in the area of gender training and legal support	a) Project planners / evaluators should inform themselves of all related initiatives in the field and coordinate with other actors in order to avoid redundancy and overlap.

Rights-based analysis

Trust Fund projects are formulated from a human rights perspective. For example, advocacy is an essential component of rights-based programming (Gosling, 03). Thus a human rights approach should be integrated into the problem analysis and the subsequent development of project hypothesis and strategy.

Definition Box 8

Rights-based analysis:
uses a conceptual framework that is normatively based on international human rights standards and operationally directed to promoting and protecting human rights.

A rights-based approach to project design implies a strong process orientation and emphasis on power relations and empowerment (Marcelino, 2005).

For impact evaluation, this implies:

- measuring changes in power relations through process monitoring;
- flexible evaluation design;
- the use of participatory methods at all phases;
- downward accountability;
- increased complexity in attributing impact;
- reliance on qualitative methods;
- process orientation to monitor staff attitudes and relationships with stakeholders;
- greater emphasis on reflective practice (Picard, 2003).

An important related tool is benefits-harms analysis for examining the context of human rights (political, security, economic, social and cultural) within which the project tries to maximize benefit and minimize harm (Caldwell, 2002).[31]

Benefits-harms analysis can be used as a project design tool that is particularly well-suited to analyze unintended effects in actually or potentially conflicted situations where change may involve risks as well as benefits for target beneficiaries.[32] Using a benefits-harms analysis, project planners in the human rights field are encouraged to examine three categories of rights and impacts: political rights, security rights and economic, social and cultural rights.

Applied to the example used in the previous sections, the considerations that flow from a benefits-harm analysis might be as follows.

> *The project aims to empower women to obtain training and advice to advocate for their own rights. Both increased claims by women and increased advocacy initiatives can incur risks and unintended consequences with regard to the three categories of rights. The political rights of women may become compromised through political infiltration or opportunism, or repressed because they develop a voice to be heard. Personal security*

[31] Rights-based approach: a conceptual framework for the process of human development that is normatively based on international human rights standards and operationally directed to promoting and protecting human rights (UNHCR website, no date).

[32] Benefits-harms analysis has been used in the medical field (oncology) and by organizations that subscribe to a human rights approach such as CARE International and the Department for International Development (DFID). The version referred to here has been developed by CARE. For guidelines to the full use of the tools, see "Benefits-Harms Handbook, CARE International, 2001.

rights of women can be threatened within their households, families and communities when their claims run counter to, or override, traditional claims. The social rights of children can become compromised when their mothers, having limited time, become involved in other activities and leave the care of small children to their school-aged daughters.

Checklist for project design

The guidelines above, applicable to new projects, have almost never been applied to existing projects, and certainly not to our sample of projects. We recommend that, to the extent possible, some of the recommended steps be undertaken retroactively with project staff to establish greater clarity with regard to objectives and indicators and to improve the pre conditions and feasibility of an impact evaluation. Aiming for a sounder evaluation basis, and given a prevailing fuzziness around objectives and indicators, the following checklist of questions should be applied prior to the design of an impact evaluation.

- What are the behavioral or institutional change objectives?
- What exactly is expected to be different as a result of the intervention?
- Is this outcome measurable?
- If yes, how is it to be observed and measured?
- Who are the direct beneficiaries and who are the ultimate beneficiaries?
- How do we draw samples of the direct beneficiaries?
- Is it feasible to take a comparison sample of direct beneficiaries?
- Does it make sense to take a sample of the ultimate beneficiaries?
- What are the assumptions about the larger environment under which the project is feasible?
- What are the factors that might influence the outcome?

It is rare to find a project manager that has addressed and tried to respond to these questions. Thus the evaluator is encouraged to go through the recommended steps with staff, even if post facto, in order to: a) gain sufficient clarification for decisions concerning the most appropriate evaluation design, and b) build planning and evaluation capacity in project staff.

3 The collection of data

We distinguish first between the collection of primary and secondary data. Primary data refers to new data, while secondary data refers to existing data such as records collected for monitoring purposes, government reports and previous studies. The advantage of secondary data is that it is far less costly than the primary data. The disadvantage is that it is collected for other purposes and rarely fits the specific data needs of the project evaluation in question.

Review of secondary data

Secondary data can come in many forms: national surveys, national policy documents, government statistics, monograms and special studies, local reports and official records, project documents and prior evaluation reports.

Secondary data is used primarily a) for a review of contextual information; b) for the systematic collection of specific data related to certain indicators; c) as a basis for sampling decisions and the establishment of comparison groups. It can also become a source for "impact monitoring" for low-cost projects having communication and advocacy objectives.

Since secondary data is collected with non-evaluation objectives in mind, the utility of secondary data needs to be assessed in terms of content, level of disaggregation, quality of data, and accessibility.

Sources of secondary data:

- Large scale national surveys: the best known national surveys conducted across a sizeable number of developing countries include the National Census, Demographic and Health Survey (DHS) by USAID/Macro International, the Social Investment Fund (SIF) Survey, the Living Standards Measurement Study (LSMS) by the World Bank, and the Multiple Indicator Cluster Survey (MICS) by UNICEF.

- Government reports and policy documents: Various ministries (health, education, etc.) issue reports, though the quality needs to be carefully assessed, especially where they are based on questionable projections of the census data (areas of high levels of migration, post conflict areas, etc.). Included are policies with regard to family law, labor law, human rights law, the Convention on the Elimination of Discrimination against Women (CEDAW), Domestic Violence Act, etc.

- Monograms and special studies: The larger international NGOs often commission special reports and assessments. While they usually address geographically specific regions, they provide useful insights into contextual issues at the broader level as well. Unfortunately, these are not always well circulated within the local NGO community.

- Local reports and official records: local government agencies (health, education, judicial system, law enforcement agencies) normally maintain records and issue routine reports. Here, too, the quality and continuity of data collection may suffer from lack of funds, staff absences, etc.

- Project or organizational documents and prior evaluations are often the point of departure for the review of secondary data.

Collection of primary data

Guided by the selection of outcome indicators, the development of data collection tools needs to consider quantitative and/or qualitative options. Quantitative evaluation methods involve the use of numerical measurement (amounts, rates, proportions) and data analysis based on statistical methods. Qualitative methods are primarily descriptive and interpretive, based on narratives, direct quotations, observations, diagrams, etc. In recent years, the combination of both approaches has become widespread practice in evaluation research.

Within each of these approaches, we recommend the use of relatively conventional methods that are already somewhat familiar to field personnel: a simple survey questionnaire for quantitative data collection and the semi-structured interview, group and/or individual, for the

collection of qualitative data. Other methods are available, but it is always underestimated how long it takes to teach good reliable practice in less well known techniques. In organizational environments, many of the better educated staff has been exposed to data collection based on interview methodologies.

Quantitative methods for primary data collection

Quantitative methods are useful for gathering information from large numbers of respondents that are representative of some population group of interest through the use of probability sampling. For TF projects, the use of such methods is recommended wherever feasible because a) of the importance of generalizing TF projects to other settings, and b) the interest in the more rigorous forms of impact evaluation.

The quantitative technique of data collection that lends itself well enough to the evaluation of most Trust Fund projects is the systematic population-based survey using a structured questionnaire. For third world settings with high cultural diversity and complexity, we prefer the interviewer-administered questionnaire. And given the prevalence of change outcomes in the area of awareness, knowledge and capacity, we further recommend survey instruments that are fashioned after the model of so-called "Knowledge, Attitude and Practice" (KAP) Surveys. Most TF projects having outcomes in the area of awareness/knowledge and capacity are appropriate candidates for this type of survey.

KAP surveys have been used in developing countries to obtain information about people's perceptions and practices, usually in the field of fertility and birth control (Warwick, 93). As a quantitative approach, it lends itself well to inquiries on knowledge, attitudes and practices in other domains such as VAW. One criticism of this type of survey is the concern that verbal statements do not necessarily reflect actual behavior, practiced or intended (Warwick, 93). We therefore recommend that questions addressing specifically the gap between knowing and doing, and that the responses related to practice be triangulated via qualitative inquiries and monitoring techniques.

A questionnaire should be constructed to facilitate the responses from the respondent as well as the social interaction between respondent and interviewer. This aspect will be important for the question domain used by TF evaluations. For TF projects, it is highly recommended to conduct at least part of the qualitative inquiries before finalizing the questionnaire.

Where the content of the questionnaire is concerned, the following general rules should be followed:

- Work as a team with a carefully selected mix of participants to discuss, develop/revise the content of the questionnaire
- Be focused and parsimonious with the content – the questionnaire should have no more than 50 questions.
- Indicate always whether only one answer applies, or whether several responses are possible.
- Introduce each thematic section so that the respondent is prepared for the subject matter.
- Use open-ended questions sparingly
- Some general rules on the wording of questions should be observed:
  - Be concise and unambiguous
  - Avoid double questions

- Avoid questions involving negatives
- Ask for precise answers
- Avoid leading questions (Alreck & Settle, 1985).

In general:

- Keep the questionnaire simple.
- Use a basic draft that has already been tested elsewhere – adapting the content to the specific project needs.
- Use a format that is already familiar in the area, for example the DHS format.

Qualitative methods for primary data collection

The focus of qualitative inquiry is to understand reality as it is perceived by the persons being studied and to understand the meaning of activities, behaviors or outcomes within the cultural context in which they are being considered (Valadez and Bamberger 1994). The benefits of qualitative assessments are that they are flexible, can be specifically tailored to the needs of the evaluation using open-ended approaches, can be carried out quickly using relatively rapid techniques, and can greatly enhance the findings of an impact evaluation through providing a better understanding of stakeholders' perceptions and priorities and the conditions and processes that may have affected program impact.

Because measuring the counterfactual (through control groups) is at the core of impact analysis techniques, qualitative methods have generally been used as a complement to quantitative evaluation techniques, providing the depth of understanding that quantitative techniques often lack. The qualitative approach uses relatively open-ended methods during design, collection of data, and analysis. Among the methodologies used in qualitative impact assessments are the techniques developed for rapid rural assessment (RRA), which rely on participants' knowledge of the conditions surrounding the project or program being evaluated.

Among the drawbacks of qualitative techniques is the lack of statistical robustness, given mainly small sample sizes, which makes it difficult to generalize to a larger, representative population. Also, the validity and reliability of qualitative data are highly dependent on the methodological skill, sensitivity, and training of the evaluator and his or her team. If field staff is not sensitive to specific social and cultural norms and practices, and nonverbal messages, the data collected may be misinterpreted. The analysis of qualitative data, coming in the form of unstructured narrative, can be extremely time-consuming, if adequately and carefully done. And finally, without a comparison group, it is impossible to determine the counterfactual and thus causality of project impact in the conventional sense (Mohr, 1995).[33]

There are many qualitative techniques available, described widely in the evaluation literature and on the internet (see links to some web based resources at the end of the guidelines). The simplest and most appropriate for our purposes, given available resources in the field and the fact that such methods are already widely used, are the semi-structured interview (individual or group) and various techniques that draw upon the PRA (participatory rural assessment) tools.[34]

[33] This implies that qualitative methods cannot be used on a stand-alone basis for impact evaluation. One notable exception would be the modus operandi method described by Scriven (Scriven, 1991). This method, however, relies on deductive skills akin to those of a forensic analyst, and therefore corresponds poorly to the constraints noted earlier.

[34] PRA techniques are widely known and many of the major international NGOs have provided training opportunities in PRA.

The semi-structured interview is a method for guiding an open-ended interview process with the intent to capture how an individual or a group of individuals think about a certain subject. It can also be combined with more visual techniques of the qualitative encounter, each time facilitating open and uncensored responses from respondents, whether they are single individuals or in groups. Responses are facilitated through a probing procedure that encourages an in-depth understanding of the respondents' perceptions and of underlying themes. While the interview protocol ensures uniformity of pre-selected themes and topics, each interview may deliver different responses based on the results of probing.

If done well, the semi-structured interview (SSI) takes time, in particular for training and the interpretation and analysis of the data. Participatory methods, such as Participatory Rural Assessment, also require time if they are to deliver good quality data.[35] Careful and disciplined management of these processes, and an optimal division of labor, can reduce time and ensure quality.

Data collection team:
The role of the interviewer/facilitator is particularly critical for the reliability and validity of data collected in this way. Because the facilitator must be fully attentive to verbal and other signals, we recommend that the data be collected by a team of two, where the facilitator is accompanied by a note-taker that can dedicate him/herself fully to the recording of the information. The use of a tape recorder is found recommended. We have found that it is best to train a note taker, providing a well structured template for recording responses. The use of tape recorders adds opportunities for disruption due to technical issues in environments that are prone to malfunction and supply ruptures (batteries, for example), it promotes a false degree of confidence as to the clarity of what is being recorded, it does not reveal the identity of the speaker, and it takes an enormous amount of time to exploit fully. A well-trained note taker, working with a well designed data entry format, can reduce the subsequent time needed for interpretation and analysis. Working as a team, the facilitator and note-taker, begin the process of reflection immediately after the interview, filling in gaps and asking for clarifications.

Interview protocol:
Whether conducted with individuals or with groups, the SSI is guided by a thematic outline that guides the interview through a series of open-ended questions and probes, facilitating open responses in the respondents' own words. The instrument used should be carefully developed with two objectives in mind: guide the recording of open-ended responses and facilitate the subsequent processing of the data. As with the questionnaire, the first section should identify the geographic location, and the type of group or individual. The second section can provide certain critical characteristics of the respondents that help in the interpretation of the responses. Training should focus on the clarification of critical conceptual categories, on the art of probing in an open-ended interview environment, and on the collaboration between facilitator and note-taker during the interview process as well as immediately afterwards with follow-up discussion, verification and triangulation.

Mixed Methods Approach

Mixed methods research is defined as "research in which the investigator collects and analyzes data, integrates the findings, and draws inferences using both qualitative and quantitative approaches or methods in a single study or program of inquiry." (*Journal of Mixed Methods*

[35] The traditional application of this method required a minimum of 3-4 days within one selected community.

Research, 2005). The mixed-method approach has in recent years found wide acceptance within the evaluation community, indicative of an ongoing trend toward pragmatism in evaluation practice (McConney et al, 2002) that is evidenced further within the WB in its preoccupation with the evaluation issues surrounding Community Driven Development (CDD) projects (discussed at 2005 Conference).

As has been noted already, differences between quantitative and qualitative methods generally involve trade-offs between breadth and depth (Patton, 2002). Both approaches offer strengths and weaknesses that can be improved or mitigated when they are combined. Thus, for example, qualitative methods can help to fill in contextual information that compensate for weaknesses in depth of quantitative techniques.

The combination of quantitative and qualitative approaches will provide a more complete, finer and broader understanding of project impact than either approach can alone (Rao & Woolcock. 2003; Wassenich & Whiteside, 2004). It is particularly well suited to project evaluation involving contextual complexity, strongly process-oriented objectives, projects that are demand-driven (limiting evaluator control), and projects faced with budget and time constraints because they permit the triangulation of information collected from smaller samples (Bamberger, 2004).

Even where causality-focused rigor is appropriate and the collection of quantitative data is feasible, the use of mixed methods is strongly recommended to address a more complete set of questions and examine the extent to which these results can be generalized to other populations. However, the most often mentioned benefit of using a mixed method approach is triangulation or the use of different sources and methods to strengthen the validity and reliability of evaluation data (Frechtling, 97).

For TF projects, the mixed methods approach will not only provide information on which projects are most effective in addressing GBV, but are more likely to explain why certain approaches are more effective. The use of the mixed methods approach influences the evaluation methodology at all levels: sampling, data collection, and data analysis. This means that at each step of the evaluation process, the evaluator determines the most appropriate quantitative and qualitative procedures, and the best integration between the two approaches.

Various options are possible for the integration of quantitative and qualitative procedures in a mixed methods evaluation. Rao (2004) differentiates between three options for the integration of methods: Parallel approaches deploy quantitative and qualitative research teams simultaneously, but bring them together during the analysis phase. The sequential option deploys one team at a time, in any order. The iterative option permits the researcher to deepen his or her understanding through on ongoing dialogue using both qualitative and quantitative procedures as needed.

Process-to-outcome monitoring

The effects of advocacy-related activities take time that is difficult to control from within a given project. The same is true of the effects of communicating new knowledge in the public sphere.

Low-cost, short term projects with advocacy and communication objectives where conventional impact assessment is not feasible may consider an outcome monitoring system. The system is intended to link outputs with individual outcomes, by monitoring, for example, the work plans of workshop participants combined with (possibly) change indicators at the level of relevant

external agencies (the judicial system, police intake, service delivery, media, government policy statements, citations, etc.). The links between outputs and outcomes can be substantiated with individual key informant interviews, or, in some cases, group interviews. The system thus combines output monitoring with tracking techniques (checklists, tally sheets) and qualitative interview methods to help understand the link between output and desired outcome. The advantage of an outcome monitoring system is that it is relatively easy to extend beyond the lifetime of the project funding period.

Table 5: Proposed indicators and data collection methods for impact and outcome objectives					
	Change objective	**Proposed outcome indicators**	**Data sources**	**Data collection methods**	**Sampling**
Impact	Increased awareness / knowledge / activities with regard to VAW Increased empowerment among women at risk	<u>Awareness / knowledge related to VAW</u>: - Existence / occurrence / spread of problem - Laws and policies - Prevention - Service options <u>Practice</u>: - Advocacy-related activities - Application / use of paralegal counseling - Networking - Participation in educational events - Number of clients having brought their case before court - Number of students / plantation workers reporting VAW incidents - Number of self-support initiatives - Number of public statements by VAW victims - Membership in VAW-related support groups and networks - Participation in decision-making structures at local level	- Survey data - Semi-structured interview data - Case load files - Police/court records - Media - Tally sheets - Semi-structured interview data	- KAP survey - Semi-structured interviews with individuals and groups to obtain contextual and process-related information - Process-to-outcome monitoring including key informant interviews	Rural women in selected regional centers High school students Police officers Legal service providers Teachers / principals Union officials Control groups: a) Randomized possible provided it is planned in advance; b) non randomized for post intervention evaluation, drawn from nonintervention sites such as schools, plantations, regional centers, etc.
Outcome	Policy and/or law reform & implementation	- Law/policy reform - Number of cases brought before court by victims of VAW - Number of rulings in support of female plaintiffs - Number of sexual abuse cases reported in schools - Number of initiatives by Mercocities municipal governments to fight VAW - Existence of regional strategy based on an integration of local municipality actions - Number of citations of research results in public documents / statements by decision & policy makers. - Number of recommendations acknowledged /	- Key informants / legal documentation / press releases - Court records / reports - School records / interview data - Municipal records - Strategy document - Number of programs aimed at reducing VAW - National committees on VAW established - Monitoring systems established - Ombudsman system established	- Process-to-outcome monitoring - Key informant interviews - Process-to-	Selected informants and informant groups All available records and press releases within one or two certain periods

Table 5: Proposed indicators and data collection methods for impact and outcome objectives					
	Change objective	**Proposed outcome indicators**	**Data sources**	**Data collection methods**	**Sampling**
Outcome	Increased knowledge through research and documentation	implemented - Policies / laws making reference to the findings / analysis / documentation - Number of citations on the internet - Number of human rights documents translated into local language	- Government reports / policy documents - Media releases - Internet - Human rights documents	outcome monitoring / individual interviews with clients	
	Increased capacity (advocacy / counseling / service delivery / networking)	- Number of clients counseled by trained community-based paralegal counselors - Applications of paralegal skills - Number of advocacy initiatives - Increased counseling activities in high schools - Existence of interstate networks for tribal women - Existence of national networks with anti VAW agenda - Number of lobbyist in support of measures against VAW - Number of referrals of VAW victims or plaintiffs - Number of requests for training related to gender, VAW-related advocacy or service delivery - VAW caseload of social service agencies - University courses on gender relations - High schools having gender relations included in curriculum - Data bases on VAW (domestic violence, trafficking, sexual harassment in schools / workplace) - Number of youth groups mobilized - Number of youth participating in peer counseling program - Number of radio/tv broadcasts on partner relations among youth - Number of parents participating in parent education programs - Number of calls per month received by hotlines - Number of calls received immediately after the crime - Number of inquiries per month received by informational lines - Number of web site visitors - Number of public debates on VAW issues	- Individual action plans - Court records - School records / interviews - Media coverage - Monitoring of social networks	- Process-to-outcome monitoring - Key informant interviews	

4 Sampling

A sample is a subset from a larger population (of people, things, events, anything) and sampling involves the strategy of selecting the subset such that it will accurately represent the patterns of the target population at large.

Definition Box 9

Sampling:
A process whereby a set of individuals or items is selected from a given population so as to provide reliable estimates of certain properties or characteristics of that population.

Sampling is applied when it is too costly or otherwise impractical to obtain information from the total population.

Not all of the TF projects justify sampling. In some cases, we are dealing with relatively limited output where it is feasible to conduct inquiries on the total population. Such is the case for projects that seek to influence legislation through the dissemination of information. The desired effect resides in the pieces of legislation and policy documents in a given subject area that have been influenced by the information. In this case, the total number of such documents is likely to be small enough to take the entire set within the appropriate time frame.

Sampling strategies are different for quantitative and qualitative approaches. "Perhaps nothing better captures the difference between quantitative and qualitative methods than the different logics that undergird sampling approaches". While quantitative samples focus on larger samples selected randomly to enhance generalization, qualitative inquiries focus typically on small samples selected purposefully to enhance understanding (Patton, 2002). Quantitative methods derive their power from representative samples that permit inferences about a larger population. Thus the size of such samples needs to be sufficiently large, in accordance to probabilistic estimates and available resources. Qualitative research is typically based on a reasoned selection of respondents or participants, emphasizing depth and descriptive detail of information rather than breadth of coverage. Thus in applying a mixed-methods approach, different sampling strategies will need to be developed, guided by the key evaluation questions in each case.

Probability sampling

For the detection of effect as a result of intervention, probability sampling should be the first choice unless it is feasible to conduct the inquiry on the entire population. The two key advantages of probability sampling for the purposes of impact evaluation is a) the ability to generalize to a broader population (external validity) and b) the ability to determine net effects (internal validity).

Probability sampling can be done in a number of ways that ensure selection by chance. We recommend two of these:

a) Simple random sample: this procedure selects randomly units from a given population. The procedure itself requires a) an enumerated list of units and b) a random number generator, usually provided by a computer program (SPSS, SAS, EPI Info and others)[36] or by hardcopy tables of random numbers.

[36] Excel has a random number generator that has been found to be less reliable than those of SPSS and SAS).

b) Systematic random sample: this is the preferred method for field situations for a number of practical reasons. First, it is less laborious when technical resources such as computerized lists are limited; second, it is more transparent when working in a participatory setting; third, it lends itself better as a teaching tool when capacity building is integrated into the evaluation process; fourth, it facilitates the replacement of missing units in the field (see also discussion by Cochran, 1976).

The difference between the two techniques depends upon context, but according to Sudman "the practical answer is that, in those cases for which simple random sampling is appropriate, simple random samples and systematic samples will be about the same except in very unusual situations and periodicities" (Sudman, 1976; Cochran, 1976).

a) Stratified random sample: This method of sampling does not preclude either of the above techniques, but refers to a strategy whereby adequate samples are ensured from minority subgroups of particular interest within a given population. The subpopulations are known as strata. Thus if the aim were to obtain information from a particular ethnic group that is generally in the minority, one would sample differentially from this subgroup to obtain an adequate number for analysis (which might be a larger proportion than from the other group). Such subgroups will be considered internally homogeneous in relation to the total population. Homogeneity within each subgroup facilitates reliable estimates from relatively small samples (Cochran, 1976).

b) Cluster samples: Cluster sampling is useful when the population is found in geographic or social clusters such as villages, urban blocks, schools, social groups. It involves the selection of clusters of observations from the population and then selects all or a random sample of elements within each cluster unit. Cluster sampling reduces cost in the field through greater ease of logistics (reduced travel costs, for example) and increases efficiency and quality through better supervision of the interviewers.

Purposive sampling

As Patton notes, what would be "bias" in statistical sampling, and therefore weakness, becomes intended focus in qualitative sampling (Patton, 2002). The intent is to select information-rich instances or cases for the purpose of in-depth information rather than a representative cross-section of a broader population. There are several strategies for purposive or reasoned sampling.[37]

- Maximum variation sampling: seeks a maximum of diversity in the selected respondents.
- Extreme case sampling: searches out extreme cases on a continuum of interest, for example, the severest cases of a phenomenon such as domestic violence.
- Intensity sampling: looks for instances that have experienced the phenomenon of interest intensely, without being it either end of the extremes.
- Homogeneous samples: are appropriate when the research focus is on a particular subgroup. The researcher will seek out members of this subgroup, such as, for example, teenage or unmarried mothers.
- Typical case sampling: searches out typical representatives of a given situation, location or phenomenon.

[37] The following examples are drawn from Patton's authoritative book on qualitative research & evaluation methods (2002).

- Critical case sampling: where resources might limit the evaluation to one single site, the researcher would look for the critical case of which one might say: "if it happens here, it could happen anywhere".
- Snowball or chain sampling: uses well-situated people or key informants as a source for selecting other informants. For example, a midwife is often a good starting point from which to select women respondents within a community.
- Criterion sampling: selects cases that meet some predetermined criterion of interest or importance.
- Opportunistic or emergent sampling: selects cases as new opportunities arise or cases of interest present themselves.
- Stratified purposeful sampling: selects different homogeneous subgroups within a larger sample. For example, poor versus middle-income women, or young versus older women, separately within a larger community of women.
- Purposeful random sampling: selects randomly participants for interviews, group discussions or case studies. This should be distinguished from a representative random sample. The purpose is credibility, not representativeness. (Note: the selection of the 15 TF projects could be considered a purposeful random sample).
- Convenience sampling: probably the most common and the least desirable method for selecting a sample. Convenience should be the last factor taken into account (Patton, 2002).

Comparison Groups

For impact evaluation involving the determination of causality and net effect, the randomized or matched selection of a control group is considered fundamental practice by most leading evaluation researchers. In most cases, such a procedure requires planning beforehand, at a time when targeting procedures and maintaining non-intervention comparison groups can be discussed and cleared with implementing agencies. Attempts to construct comparison groups after project implementation becomes usually more complex and costly, especially for projects without baseline data.

Bamberger, addressing situations of real-world constraints, has explored options for addressing such constraints, though concedes that for some (or even many) projects, impact evaluation may not be feasible at all or no longer feasible at some point (Bamberger, 2005). This would appear to be the case for the selected Trust Fund projects. We can, however, propose sampling options that might have been considered at the outset of these interventions. We will also consider "good enough" sampling strategies for evaluations of project effect with reasonable (though approximate) estimations of net impact.

Proposed options for selected Trust Fund projects

The target groups identified during the project design phase become the basis for this fundamentally critical component of evaluation design. Indeed, the quality of impact evaluation depends on a clear and precise definition of the target group and in the establishment of clear boundaries that identify who is included and for what reason, and who is not. As Rossi notes, "Although target definitions may be easy to write, it is often difficult to employ such definitions in the more precise work of needs assessment and program design. There are few human and social problems that can be easily and convincingly described in terms of simple, unambiguous characteristics of the individuals experiencing that problem." (p.139)

When identifying target groups, a common tendency is to make those boundaries too ambiguous, broad and overinclusive. Discussing the sampling frame, the sampling selection

procedure and the possibility of control groups encourages a more concrete and practical identification of the target groups.

Example one:
Key change objective: awareness raising
Project duration: 1 year

The direct target group consists of young people (15-30 years) in four selected communities in Suriname. The sampling process should focus on representative samples of young men and women in the four pilot communities, matched by representative samples in equivalent communities (targeted for future intervention) that are not included in the current initiative. In order to facilitate the selection of this sample, the enrollment lists of the local secondary schools could be used. This would restrict the sample to students aged about 15-21 for the quantitative study. A qualitative sample of young people within this age group that are not attending school could be selected from community youth groups.

Example two:
Key change objective: advocacy / capacity building
Project duration: 1 year

The direct target groups for this intervention are teachers, union officials, female students and plantation workers in selected rural pilot areas. The intervention consists of training workshops. The quantitative sample should be selected from the complete list of workshop participants, possibly using a stratified approach to ensure an adequate sample from each of the subgroups. The qualitative samples should select key informants from the same subgroups among non-participants a) within the pilot areas, b) within matched areas outside the pilot zone. At the broader level, evidence of impact at the policy level could be monitored and recorded on an ongoing basis. Key movers of new policies or policy revisions should be targeted for semi structured informant interviews.

Example three:
Key change objective: capacity building
Project duration: 1 year

The direct target group of the intervention at the effects level (change in behavior or practice) is community-based persons from a fairly broad range of professional backgrounds that are trained to provide legal support to poor women in the 10 rural provinces of Ethiopia. With about 30 trainees per province, and given their educational level, a self-administered KAP questionnaire could be administered to all trainees. This could be done for some time (2-3 years) after the training has been completed, since recall of actions taken on the basis of this training should not be too difficult. Since this project was planning a second round of training, the best option for selecting a control group would have been to take a “pipeline approach”, where the participants for subsequent trainings receive the same self-administered questionnaire prior to their training.

Specific sampling issues and procedures

Sample size:

TF projects, being relatively short-term, are primarily oriented to the achievement of effects at the level of attitude, knowledge and/or practice. They often work directly with intermediary

groups through whom the ultimate beneficiaries (actual or potential victims of VAW) stand to benefit. The problem of calculating sample size should therefore be less daunting, since they rarely require large population-based samples.

In the absence of the services of a costly statistician, the following practical recommendations may simplify the determination of the size of quantitative samples for TF projects.

- Where direct target groups are of the intermediary kind, such as participants of training events, it is often possible to take the entire participant population based on the following justification: a) within the limits of a one-year project, the number of participants is unlikely to exceed 400; b) this group can be provided with a self-administered survey questionnaire, saving interviewer cost and time.

- Where target groups are a subset of the population at large, such as teachers, law enforcement officials or social service agents, cluster sampling based on schools, judicial precincts, social service catchment areas or the like should be considered.

- Where targets are institutions (for example, the Mercocity Project, #15 in the project summary in Appendix 1), the entire population of institutions should be taken for "outcome monitoring" – that is, the tracking of their initiatives in the area of interest. Using purposive sampling, the most appropriate representatives should be identified for key informant interviews. However, self-administered questionnaires should be avoided. They do not usually work well in bureaucratic settings.

5 Data Processing and Analysis

Facts are terrible things if left sprawling and unattended
(Norman Cousins, 1981, quoted by Weiss, 1998).

Good data analysis is an inherently iterative process and similar basic skills of logic, inquisitiveness and analytical thinking are required at all levels. For the processing and analysis of relatively simple data sets, it is recommended that the analyst oversee the entire process. The mechanistic compartmentalization of these tasks, frequently encountered in developing countries, is not recommended.

Data Processing

By data processing we refer to the entry and cleaning of data, and it's preparation for analysis. Quantitative data will require at the very least lengthy tabulations that usually exceed the capacity of simple calculators working from raw data input sheets.

The questionnaire should become the basis for a data entry template that facilitates entry into a computerized data matrix. This process is fortunately well facilitated by off-the-shelf survey software programs that perform most functions related to the processing, cleaning and preliminary analysis of data. [38] In some cases, Access and Excel have been programmed for these purposes, but their functionality is not as specifically oriented toward survey data processing.

[38] We recommend, in particular, two widely used software programs that can be downloaded free of charge: EpiInfo from the Center for Disease Control, and CSPRO from the U.S. Bureau of the Census. Their websites are provided in appendix xx.

Having the software and skills to use it is one thing, organizing the process by which data input forms go out and completed forms are verified and returned for entry and cleaning is another. What is critical is to have a solid quality control structure where the content of the data forms are checked at critical points (the first being immediately after the interview encounter), verified, coded (in the case of open-ended questions) and submitted for data entry.

Data Analysis

"The aim of analysis is to convert a mass of raw data into a coherent account" (Weiss, 1998).
The analysis of data follows a logic that should determine or influence all other phases related to the collection and analysis of primary data, from the development of the data collection instrument(s) to the presentation of results. Evaluation is an iterative process, and the analysis of information should not be a discrete and separate activity but should be integrated into all phases of the process.

Analysis plan

The development of an analysis plan early on and displayed visibly, helps to keep analysis questions at the forefront throughout the evaluation process, including the finalization of the data collection instruments. It also forces evaluators to clarify questions related to measurement and to focus evaluation activities.

An analysis plan should:

- present the key unit(s) of analysis (households, individual respondents) together with basic descriptive data (geographic location, basic characteristics such as age, sex, marital status) that will help to replicate the study. This information is often critical for matching purposes as well.
- provide carefully selected basic demographic and social background characteristics that might influence responses (e.g., education, marital status, occupational category, ethnicity) and which may have explanatory value.
- present the key response categories (for example, related to knowledge, attitude and practice) by key comparative categories (for example, geographic location or rural versus urban).
- follow explicitly the original project objective statements and indicators. Thus in the report, it is a good idea to dedicate a clearly labeled section to each objective, listing related indicators beneath it or in an introductory paragraph.

Quantitative data

Three items are usually required for this phase of evaluation:

- Simple and robust survey software
- An analysis plan
- A skilled evaluator/analyst

The first ingredient is fortunately readily available and in some cases free of charge (see web related resource list at and of guidelines). At the UNIFEM level, I would recommend investment in SPSS (Statistical Package for Social Sciences) for more flexibility and ease of analysis, and because it is so widely used and known. These software packages come with online manuals that can provide support to someone having had basic training in modern research and

computing techniques, a fair amount of intrinsic curiosity, perseverance and time if they are new to the software. National statistical departments often have staff that has been trained in data analysis and that can provide training support on a short-term contractual basis.

The third element is more difficult to secure. Good analytical skills are not always readily available. In countries with experience in the large international surveys such as the Demographic and Health Survey (DHS), or the Multiple Indicator Cluster Survey (MICS), skills for quantitative analysis are often available through contacts with national statistical institutes or regional/national evaluation associations. It is at this level that donors should ensure sufficient support to project-based staff. The preferred scenario for small and medium-sized project evaluations is that the evaluation leader and analyst be the same person. Unfortunately, one often finds the practice of compartmentalizing different skills, assigning the computer work to a statistician who delivers the output according to the requests of the evaluator. Human rights projects are far too complex for such a compartmentalization.

The analysis of quantitative data, guided by an analysis plan that has been communicated in advance, should be conducted as transparently as possible, facilitating interpretation before a general and non-technical audience.

Qualitative data

While the fundamental logic of inquiry remains the same, the analysis of qualitative data requires a different sequence of procedures. Instead of anticipating the range of responses and coding response categories in advance, qualitative analysis takes data in its raw and narrative form and organizes it in terms of critical information categories or "meaning units" (Cresswell, 2003).

The following basic steps are recommended:[39]

- Begin with a review of the raw data, jointly with the persons that collected it and other parties that might be able to shed light on questions or ambiguities when they arise.
- Organize the information into thematic categories and coded, setting aside representative or explanatory verbatim quotes for key categories. Again, it is highly recommended to do this process as a team, engaging project staff and other resource persons in the exercise. If the project design process has developed a problem tree, this document or diagram should be visually available for conceptual reference. Codes should be nested by organizing the more specific response codes by subheadings which in turn fall under a given broader thematic heading. (For example: domestic violence =100; physical/psychological/ economic/…=110/120/130/…; specific codes within 110=110-119).
- Record the results of the coding process in a "code book" kept for ongoing revision, future reference and for a possible replication of the study. Ensure that the last version is clearly identified and dated. Assign primary responsibility to one person for updating and revising the codebook (MacQueen, 98).
- Depending on the quantity of qualitative data entries per unit of analysis (key informant, groups), the coded information can be entered into a data matrix and linked to key identifying categories used for the quantitative data (e.g., geographic location).

[39] Software exists for this operation, but does little to reduce the time or conceptual labor required, and tends to distract from the participatory and organic nature of group work.

Combining qualitative and quantitative data analysis

While the analytic process is distinct for each type of data, the two sources of information can be integrated during an interpretive process that confronts the quantitative findings with qualitative profiles of broader groupings (rural versus urban settings, for example) and contextually descriptive information.

The intent is to strengthen the validity of the evaluation, especially where resources do not permit highly sophisticated studies and where the strengths of each method can compensate for the weaknesses of the other method. The principal validity issues that are likely to benefit from a mixed methods approach are:

Internal validity: refers to the extent to which the intervention delivers the expected effects. Where quantitative techniques can deliver reliable estimates of such effects, qualitative information can help to understand, for example: a) why such effects are not uniform throughout the study population; b) what implementation conditions can enhance (or minimize) effects; c) what explains the presence of unintended effects.

External validity: refers to the extent to which the intervention-to-effect relationship be generalized to other settings or populations? Quantitative sampling can ensure that the results are generalizable to the broader population within which the study population resides. But questions of generalizability to other social or cultural settings are better examined with the help of qualitative and contextual data.

6 Interpretation and Presentation of Evaluation Results

Interpretation and presentation, oral and written, are critical stepping stones to the utilization of evaluation results. They are iterative processes that should be as collaborative and participative as possible in order to get maximum buy-in. All stakeholders, including the field teams, should have an opportunity to provide input prior to the finalization of the evaluation report.

Preliminary oral presentations and validations

There are several ways of facilitating a collaborative/participative interpretation and validation process.

- Request data collection reports from each field supervisor that provide observations on the data collection process and on any issues that might influence the quality of the data. Discuss these collectively at the end of the data collection process.
- Ensure continued access to field supervisors for questions that arise while cleaning the data and during the review of the first tabulations.
- Conduct a well-structured meeting with project staff and/or other stakeholders to review preliminary results. Ensure that a note taker record all questions and comments.
- In the case of a finding that appears to be counter intuitive, organize a small special inquiry together with staff. I have found this technique to yield excellent returns in improved understanding, learning, and buy-in into the evaluation results.

Oral presentations should be clear about their specific purpose and the audience to be addressed. Two important objectives of oral presentations that precede the finalization of a written report are a) to get interpretive feedback to preliminary results, and b) to discuss the final

results and their implications for lessons learned and future action. In both cases, the opportunity for input should be maximized and feedback carefully noted and integrated into revisions, clarifications and the finalization of the written report.

The evaluation report

The final evaluation report combines the results of the entire process and organizes it so that it responds to the information needs of grantees and donors as clearly as possible. The report should be self-explanatory, providing results in summary form as well as in relevant detail.

Producing a report outline at the outset helps to structure the process of writing.
The outline should include the following components:

Executive summary: this summary draws upon the final report, and while it is presented in the beginning of the final document, it is usually the final main section to be written. The executive summary should be no longer than 2-3 pages, providing only the major highlights of the results.

Introduction: The introduction permits to remind the reader of the key purpose of the impact evaluation, of the key project objectives, and of the indicators that the evaluation set out to measure. It also provides a brief summary of the project, its implementation and funding context.

Methodology: The purpose of this section is to summarize methodological aspects of the study, providing supporting documentation such as data collection instruments in the appendices:
a) methodological design, including the use of baseline data.
b) sampling methodology, providing a table with the distribution of the sample(s) (study group and comparison group where applicable) across appropriate key categories, such as, for example, sex, geo-administrative unit, urban versus rural environments.
c) data collection instruments used, with a list of key themes addressed by each (provide copies of questionnaires and guides in the appendix)
d) process of data collection with time frame (very important) and number of field interviewers deployed (provide calendar of activities in appendix).

Results: The structure and language of this section needs to be as user-friendly as possible.
The results chapter should end in a concise conclusion that links, once again, between project objectives, indicators, measures and results.

Recommendations: The recommendations provide an opportunity to step back and situate the results in the broader context of project and donor objectives. If possible, the recommendations should be the result of the discussions held with stakeholders in the course of prior presentations.

BIBLIOGRAPHY

Alreck, Pamela L. & Settle, Robert B. *The Survey Research Handbook.* Richard D. Irwin, Inc. 1985.

Arndt, Nin A.C., Preckel, P.V. Is agricultural productivity in developing countries really shrinking? New evidence using a modified nonparametric approach. *Journal of Development Economics*, 71(2): 395-415.

Aubel, Judi. Participatory Program Evaluation Manual. Child Survival Technical Support Project and Catholic Relief Services. 1999

Baker, Judy L. Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners. The World Bank. 2000.

Bamberger, Michael. Designing quality impact evaluations under budget, time and data constraints. Presentation to Poverty Analysis, M&E Thematic Group on July 28, 2005

Bamberger, Michael. The Evaluation of International Development Programs: A View from the Front. *American Journal of Evaluation*. Vol.21, Number 1, Winter 2000.

Bamberger, Michael et al. Shoestring Evaluation: Designing Impact Evaluations under Budget, Time and Data Constraints. *The American Journal of Evaluation*. Vol.25, No.1, Spring 2004.

Barton, Tom. Guidelines to Monitoring and Evaluation: How are we doing? CARE International in Uganda. 1997.

Bogat, G.Anne, Lenvendosky, Alytia A. & Davidson, William S.II. Understanding the Intergenerational Transmission of Violence from Pregnancy Through the First Year of Life, Final Project Report. US Department of Justice. October 2002.

Bott, Sarah, Morrison, Andrew and Ellsberg, Mary. Preventing and responding to gender-based violence in middle and low income countries: a global review and analysis. World Bank Policy Research Working Paper3618, June 2005.

Caldwell, Richard. Project Design Handbook. Tango International for CARE International. 2002

CARE International. Benefits-Harms Handbook. Nairobi, Kenya 2001

Chapman, Jennifer & Wameyo, Amboka. Monitoring and Evaluating Advocacy: A Scoping Study. Paper written for ActionAid. January 2001.

Chase, Robert. Impact Evaluations: Practical Approaches to a Tough Analytic Problem. Presentation to Poverty Analysis, Presentation at M&E Thematic Group, March 15, 2005.

Chen, Huey-Tsyh. *Theory-Driven Evaluations*. Sage Publications, Inc., 1990.

Cochran, William G. *Sampling Techniques*. 3rd ed. John Wiley & Sons, 1976.

Cook, Thomas D. and Campbell, Donald T. Quasi-Experimentation: Design & Analysis Issues for Field Settings, Houghton Mifflin Company, Boston, 1979.

Cresswell, John.W. *Qualitative Inquiry and Research Design: Choosing Among Five Traditions*. Sage Publications. 1998.

Cresswell, John.W. *Research Design: Qualitative, Quantitative and Mixed Methods Approaches 2nd ed.* Sage Publications, Inc., 2003.

Davies, Rick. Network Perspectives in the Evaluation of Development Interventions: More Than a Metaphor. EDAIS Conference, November 2003.

De Toma, Costanza and Gosling, Louisa. Advocacy Toolkit: a collection of tools to help plan, implement, monitor and evaluate advocacy. Save the Children. March 2005.

Den Heyer, Molly. The Temporal Logic Model: A Concept Paper. IDRC Evaluation Unit. July 2001.

Department for International Development. Monitoring and Evaluating Information and Communication for Development (ICD) Programmes Guidelines. March 2005

Donaldson, Stewart I. and Christie, Christina A. The 2004 Claremont Debate: Libsey vs. Scriven. *Journal of Multi Disciplinary Evaluation*. Number 3, October 2005.

Duflo, Esther. *Scaling Up and Evaluation*. Paper presented at the Annual World Bank Conference on Development Economics 2004.

Duflo, Esther and Kremer, Michael Use of Randomization in the Evaluation of Development Effectiveness. Paper prepared for the World Bank Operations Evaluation Department (OED), Conference on Evaluation and Development Effectiveness. July 2003.

Dugger, Celia. World Bank Challenged: Are Poor Really Helped? *New York Times*. July 28, 2004.

Earl, Sarah, Carden, Fred and Smutylo, Terry. Outcome Mapping: Building Learning and Reflection into Development Programs. IDRC. 2001.

Frechtiling, Joy; Sharp, Laure and Westat, Inc., Editors. *User-friendly Handbook for Mixed Method Evaluations.* National Science Foundation.Division of Research, Evaluation and Communication. 1997.

Frechtling, Joy. *The 2002 User Friendly Handbook for Project Evaluation*. National Science Foundation, Division of Research, Evaluation and Communication. 2002.

Fried, Susana T. UNIFEM's EVAW Assessment: Summary of the Regional Scans. UNIFEM. 2001.

Gennetian, Lisa A., Bos, Johannes M. & Morris, Pamela A. Using Instrumental Variables Analysis to Learn More from Social Policy Experiments. MDRC. October 2002.

Gennetian, Lisa A. US Welfare Policy Effects on Domestic Abuse: Evidence from Random Assignment Studies. Presentation at the Development Implications of Gender-Based Violence Workshop. World Bank. November 2004.

Germann, Dorsi and Gohl, Eberhard. PIM Booklet 2: NGO-Based Impact Monitoring. GTZ-GATE. 1996.

Gosling, Louisa. Monitoring and Evaluating Advocacy Programmes. *Aid Workers Exchange*. October 2003.

Hansen, Hanne Foss. Choosing Evaluation Models: A Discussion on Evaluation Design. *Evaluation*. Vol. 11(4):447-462. Sage Publications. 2005.

Ingram, Joseph. Statement to the Sub-Commission on the Promotion and Protection of Human Rights-57th Session. July 25-August 12, 2005.

International Fund for Agricultural Development (FAD), Office of Evaluation and Studies. *A Guide for Project M&E – Managing for Impact in Rural Development.* Rome, Italy.

Johnson, R.Burke and Onwuegbuzie, Anthony J. Mixed Methods Research: A Research Paradigm Whose Time has Come. *Educational Researcher*, Vol 33, No.7, pp.14-26. 2004.

Keogh, Erica. Developing a framework for budgeting for household surveys in developing countries. Survey Preparations in Developing Countries. 2002.
Kish, Leslie. *Survey Sampling.* John Wiley & Sons. New York. 1965.

Kishor, Sunita. Domestic violence measurement in the demographic and health surveys: the history and challenges. Expert paper presented at the expert group meeting, Geneva, Switzerland, 11-14 April, 2005.

Kishor, Sunita. Profiling Domestic Violence: A Multi-Country Study. Measure DHS+. ORC Macro. June 2004.

Koenig, Michael et al. Domestic violence in rural Uganda : evidence from a community-based study. *Bulletin of the World Health Organization*, 81 (1), 2003.

Kremer, Michael. Randomized Evaluations of Educational Programs in Developing Countries: Some Lessons. *American Economic Review* 93(2), May 2003.

Kwong, Marilyn J. et al. The Intergenerational Transmission of Relationship Violence. *Journal of Family Psychology*, 2003, Vol 17, No.3, 288-301.

Kumar, Yogesh. Monitoring and Evaluation of Advocacy Campaigns: Opportunities and Challenges. Asia Workshop on Next Generation Participatory Monitoring and Evaluation. November 2005.

Lalasz, Robert. Domestic Violence in Developing Countries: An Intergenerational Crisis. Population Reference Bureau. 2006.

Lawrenz, Frances and Huffman, Douglas. The Archipelago Approach To Mixed Method Evaluation. *The American Journal of Evaluation, Volume 23, Issue 3, Autumn 2002, Pages 331-338*

Lyon, Eleanor. Impact Evaluation of Special Session Domestic Violence: Enhanced Advocacy and Interventions. US Department of Justice. June 2005.

MacQueen, Kathleen M. et al. Codebook Development for Team-Based Qualitative Analysis. *Cultural Anthropology Methods*, 10(2): 31-36. 1998.

Malhotra, Anju, Schuler, Sdiney R. & Boender, Carol. Measuring Women's Empowerment as a Variable in International Development. Paper presented to the World Bank. June 2002.

Mansuri, Gazala & Rao, Vijayendra. Community-Based and –Driven Development: A Critical Review. The World Bank Research Observer, vol.19, no.1, 2004

Marcelino, Elena. A Rights-Based Approach to Evaluation. UNIFEM's Initial Experience. Paper presented at the Topical Session on A Rights Based Approach to Evaluation

Mason, Karen Oppenheim. Measuring Empowerment: A Social Demographer's View. Paper presented to the World Bank. February 2003.

Mayoux, Linda. "What do we Want to Know? Selecting Indicators". Paper submitted to WISE.

McConney, Andrew; Rudd, Andy; and Ayres, Robert. "Synthesizing Findings Within Mixed-method Program Evaluations". *American Journal of Evaluation,* Vol.23, No.2, 2002, pp.121-140.

McNamara, Carter. Basic Guide to Outcomes-Based Evaluation for Nonprofit Organizations with Very Limited Resources. Free Management Library. 1999.

Meager et al. Awareness, Knowledge and Exercise of Individual Employment Rights. Employment Relations Research Series No.15. Department of Trade and Industry. 2002.

Middleton, Andrew. Logical Framework Analysis: A Planning Tool for Government Agencies, International Development Organizations and Undergraduate Students.
Mohr, Lawrence B. *Impact Analysis for Program Evaluation*. Sage Publications. 1995.

Miller, Thomas I., Kobayashi, Michelle M., and Noble. Paula M. Insourcing, Not Capacity Building, a Better Model for Sustained Program Evaluation. *American Journal of Evaluation*. Vol.27, No.1, March 2006.

Morley, Elaine, Vinson, Elisa and Hatry, Harry P. Outcome Measurement in Nonprofit Organizations: Current Practices and Recommendations. Independent Sector. 2001.

Morrison, Andrew R. and Orlando, Maria Beatriz. The costs and impacts of gender-based violence in developing countries: Methodological considerations and new evidence.

Ogata, Sadako. Empowering People for Human Security. Payne Lecture delivered at Stanford University on 28 May, 2003.

Osborne, Jason W. & Costello, Anna B. Sample size and subject to item ratio in principal components analysis. Practical Assessment, Research & Evaluation, 2004, 9(11).

Patton, Michael Quinn. "The Debate about Randomized Controls as the Gold Standard in Evaluation". Presentation to the International Finance Corporation, May 2005.

Patton, Michael Quinn. Qualitative Evaluation Checklist. www.wmich.edu/evalctr/checklists. September 2003.

Patton, Michael Quinn. *Qualitative Research and Evaluation Methods*. 3rd ed. Sage Publications Inc. 2002.

Picard, Mary. Measurement and Methodological Challenges to CARE International's Rights-Based Programming. Paper submitted to the EDIAIS Conference on "New Directions in Impact Assessment for Development: Methods and Practice", Manchester, UK. November, 2003.

Rao, Vijayendra and Ibanez, Ana Maria. The Social Impact of Social Funds in Jamaica: A 'Participatory Econometrics' Analysis of Targeting, /Collective Action, and Participation in Community-Driven Development. *The Journal of Development Studies,* Vol.41, No.5, July 2005, pp.788-838.

Rao Vijayendra and Woolcock, Michael. Integrating Qualitative and Quantitative Approaches in Program Evaluation. The Impact of Economic Policy on Poverty and Income Distribution: Evaluation Techniques and Tools, World Bank and Oxford University Press, 2004.

Ravallion, Martin. Evaluating Anti-Poverty Programs. Prepared for the Handbook of Agricultural Economics Volume 4, North-Holland, edited by Robert E. Evenson and T. Paul Schultz, 2005.

Richards, Paul; Bah, Khadija & Vincent, James. Social Capital and Survival: Prospects for Community Driven Development in Post-Conflict Sierra Leone. *Social Development Papers*. World Bank. April 2004.

Roche, Chris. *Impact Assessment for Developing Agencies: Learning to Value Change*. Oxfam. 1999.

Rosen, Leora N. Violence & Victimization Research Division's Compendium of Research on Violence Against Women 1993-2005. National Institute of Justice. 2005.

Sandelowski, Margarete. Combining Qualitative and Quantitative Sampling, Data Collection, and Analysis Techniques in Mixed-Method Studies. *Research in Nursing and Health*. Vol.23. 2000: 246-255.

Schafer, William D. Replication: a design principle for field research. *Practical Assessment, Research & Evaluation*, 2001, 7(15).

Scriven, Michael. *Evaluation Thesaurus*, *4th ed*. Sage Publications. 1991.

Spindel, Cheywa, Levy, Elisa and Connor Melissa. With an End in Sight: Strategies from the UNIFEM Trust Fund to Eliminate Violence Against Women. UNIFEM. 2000.

Smulovitz, Catalina & Walton, Michael. Evaluating Empowerment. Paper presented at the World Bank. February 2003.

Sudman, Seymour. Applied Sampling. Academic Press, Inc., San Diego, CA, 1976.

Tomison, Adam M. Intergenerational Transmission of Maltreatment. *Issues in Child Abuse Prevention*, 1996, No.6.

Trevisan, Michael S. and Huang, Yi Min. Evaluability Assessment: A Primer. *Practical Assessment, Research & Evaluation*. December 2003

UNDP. Poverty Reduction and Human Rights: A Pratice Note. June 2003.

UNICEF. Domestic Violence Against Women and Girls. Innocenti Digest No.6. 2000.

UNIFEM. Making a Difference: Strategic Communications to End Violence against Women. 2003.

UNIFEM. Not a Minute More: Ending Violence Against Women. 2003.

UNIFEM. The Trust Fund in Support of Actions to Eliminate Violence Against Women: A New Strategy. 2004.

UNIFEM. With an End in Sight. 2003.

United Nations, General Assembly. The role of the United Nations Development Fund for Women in eliminating violence against women. A/RES/50/166. December 1995.

United Nations Development Fund (UNDP). Handbook on Monitoring and Evaluating for Results. United Nations Development Programme Evaluation Office. New York. 2002.

United Nations Development Fund for Women (UNIFEM). *Not a Minute More: Ending Violence Against Women*. UNIFEM, New York, 2003.

United Nations Development Fund for Women (UNIFEM). Operational Guidelines. Trust Fund in Support of Actions to Eliminate Violence Against Women. 2001.

United Nations Development Fund for Women (UNIFEM). Operational Guidelines for Cycle 9. Trust Fund in Support of Actions to Eliminate Violence Against Women. 2004.

United Nations Development Fund for Women (UNIFEM). Project files on 22 Trust Fund projects. Trust Fund archives in New York. Trust Fund in Support of Actions to Eliminate Violence Against Women. 1997-2004.

United Nations Statistics Division. Household Sample Surveys in Developing and Transition Countries. Department of Economic And Social Affairs. Studies in Methods: Series F no.96. 2005

University of Reading, Statistical Services Centre. Some Basic Ideas of Sampling. Statistical Good Practice Guidelines. November 2000.

Uphoff, Norman. Some Analytical Issues in Measuring Empowerment for the Poor, with Concern for Community and Local Governance. Paper presented to the World Bank. February 2003.

Valadez, Joseph and Bamberger, Michael. Monitoring and Evaluating Social Programs in Developing Countries: A Handbook for Policymakers, Managers and Researchers. Economic Development Institute of The World Bank, 1994.

Wassenich, Paul and Whiteside, Katherine. CDD Impact Assessments Study: Optimizing Evaluation Design Under Constraints. Social Development Papers: Community Driven Development, No.51, February 2004.

Weiss, Carol. *Evaluation: Methods for Studying Programs and Policies.* Second Editions, Prentice Hall. New Jersey. 1998.

Whiteside, Katherine; Woolcock, Michael and Briggs, Xavier de Souza. Assessing Social Development Projects: Integrating the Science of Evaluation and the Art of Practice. Poverty Impact Analysis, Monitoring and Evaluation Thematic Group
World Bank, Washington, DC, December 2004

W.K.Kellogg Foundation. Logic Model Development Guide: Using Logic Models to Bring Together Planning, Evaluation & Action. October 2000.

W.K.Kellogg Foundation. Evaluation Handbook. January 1998.

World Bank, Independent Evaluation Group (IEG). "OED and Impact Evaluation: a Discussion Note". (no date).

World Conference on Human Rights. Vienna Declaration and Programme of Action. 1993.

Yansaneh, Ibrahim S. An analysis of cost issues for surveys in developing and transition countries. Household Sample Surveys in Developing and Transition Countries. Chapter XII. United Nations Statistics Division. 2005

WEB BASED RESOURCES FOR PROJECT DESIGN AND EVALUATION

The following list of resources is organized by topic area and are readily accessible at the time of this writing.

Project design

The following link from the University of Wisconsin extension program provides useful guidelines to, and examples of, logic models.

http://www.uwex.edu/ces/pdande/evaluation/evallogicmodel.html

This link provides access to the W.K.Kellogg Foundation's Logic Model Development Guide.

http://www.wkkf.org/Pubs/Tools/Evaluation/Pub3669.pdf

The link below provides access to the so-called "Rosetta Stone of Logical Frameworks", presenting a comparison of terminologies used by different donor agencies for logical or results framework categories.

http://www.mande.co.uk/docs/Rosettastone.doc

Other guides to the use of LFA tools can be found through the following websites: www.jiscinfone.ac.uk; www.ifad.org/evaluation/guide; www.web.mit.edu.

Evaluation design

A basic guide to experimental design from the University of Reading's Statistical Services Centre.

http://www.rdg.ac.uk/ssc/publications/guides/topdoe.html

An introduction to quasi-experimental design, from Professor Trochim of Cornell University.

http://www.socialresearchmethods.net/kb/quasioth.htm

Sampling

Also from the University of Reading's Statistical Services Centre, a good basic introductory guide to sampling.

http://www.rdg.ac.uk/ssc/publications/guides/topbis.html

This site provides an introduction to probability sampling from Prof. Trochim.

http://www.socialresearchmethods.net/kb/sampprob.htm

On sampling procedures :

http://www.socialresearchmethods.net/kb/sampprob.htm

Survey methodology

The following site has several links to technical guidelines to household survey methodology and implementation.

http://gsociology.icaap.org/methods/surveys.htm

From the University of Reading's Statistical Services Centre

http://www.rdg.ac.uk/ssc/publications/guides/toppes.html

Questionnaire design

http://www.leeds.ac.uk/iss/documentation/top/top2/top2.html

Focused on the health field, this link provides access to evaluation experience documented by Johns Hopkins University, including questionnaires used in impact assessments.
http://www.jhuccp.org/research/

The link below is to the DHS multi-country study on domestic violence, which includes the questionnaire segment that was included in the DHS surveys.
http://www.measuredhs.com/pubs/pdf/OD31/OD31.pdf

Qualitative Methods

This is a very useful guide by a long-time practitioner.

Aubel, Judi. Participatory Program Evaluation Manual. 1999
http://www.idrc.ca/uploads/user-S/10504133390Participatory_Program_Evaluation_Manual.pdf

Patton, Michael Quinn. "Qualitative Evaluation Checklist", September 2003. Available online in English at http://www.wmich.edu/evalctr/checklists/qec/index.htm

International Centre for Development-oriented Research in Agriculture
Informal Interview Guidelines
http://www.icra-edu.org/objects/anglolearn/Informal_Interviews-Guidelines.pdf

Guidelines for the coding of qualitative data :

http://www.cdc.gov/hiv/software/pubs/codebook.pdf

This website provides a list of links to resources on rural appraisal methods including PRA and RRA (Rapid Rural Appraisal), as well as to resources on the conduct of formal surveys.

www.icra-edu.org/page.cfm?pageid=anglolearnmethodslinks

Survey analysis software

EpiInfo: A widely used software that serves a range of survey-related tasks including analysis is issued by the Center for Disease Control, free of charge and available at:

www.cdc.gov/epiinfo/about.htm

CSPro: The people who conduct the worldwide DHS (Demographic and Health Surveys) use an excellent software for data entry, processing and some basic analysis that is available free of charge:

www.census.gov/ipc/www/cspro/index.html

Both of these software packages come with fairly extensive documentation.

General

This site makes available a multitude of useful links on a broad range of issues related to research and evaluation. It also includes links to free software for surveys and analysis.

http://gsociology.icaap.org/methods/

This is an excellent resource site for evaluators. It is to be updated in March 2006.

http://www.resources4evaluators.info

The same site provides access to many how-to documents and tools:

http://www.resources4evaluators.info/EvaluationTextsAndDocuments.htm#Tools

This site makes available useful guidelines for various aspects of evaluation research.

http://www.rdg.ac.uk/ssc/publications/guides.html

A Manual on Outcome Mappling issued by the IDRC can be found at the following site:

http://www.idrc.ca/en/ev-9330-201-1-DO-TOPIC.html

APPENDICES

1 Summaries Of Selected Projects

1 Title: Legal training of community-based women

The objective of this project was to educate Ethiopian women in rural communities about the legal instruments available to them and the manner in which they can be used in instances of gender-related crimes and violence. With no practicing women lawyers outside of the capital, women have very limited access to information that they can use in their own defense. The project trained women to become community-based paralegals specializing in laws related to crimes against women. A 15-day training program was developed and implemented in all of Ethiopia's 10 regions outside the capital. The proposed beneficiaries are the women in Ethiopia's rural communities. The intermediate participants in the training program consist of 301 committee members of the EWLA, 259 women and 42 men from various professional backgrounds: teachers, police agents, health workers, psychologists, community workers, secretaries and government employees.

Project objectives:
- To raise the knowledge and awareness of EWLA regional committee members so that they can use the laws for their protection and to redress crimes directed against them.
- To provide women with basic legal skills so they can assist women in their community.

Principal change/outcome objectives:
Awareness/knowledge
Para-legal skills
Empowerment

Key strategies:
Education and training

The project received $64,865 from the Trust Fund for the duration of one year, with $5,000 set aside for monitoring and evaluation costs.

2 Title: Ending violence against Kenyan women

This project responds to the high incidence of rape committed in Kenyan schools by male students and teachers. Police statistics report daily incidences of defilement (and/or rape) of girls, indicating that the actual rate is much higher due to known underreporting. The project targets plantations where women provide 80% of the labor and where existing unions have no mechanisms for addressing issues related to gender-related injustice and violence. Project activities consist of training workshops and seminars conducted for government and trade union officials, teachers and women, and officals of the various institutions involved (Ministries of Labor and Education, etc.). The project also conducted empowerment workshops for female plantation workers and women in educational institutions and developed a training curriculum and manual for future activities. The overall objective of the project is "to implement training programmes for education and trade union officials that would sensitize them about gender

violence" and furthermore "The proposed activities are meant to supplement the Legal Awareness Project in areas into which we would want to extend our activities" (project proposal, June 97). Thus a primary objective is to raise awareness with an eye to influence policy and advocacy for law reform on sexual abuse and the violation of women's rights.

Project objectives:
- To conduct training seminars and workshops for the respective officials on women's rights and gender issues.
- To assist the two departments and trade unions to develop policies and advocacy for law reform on sexual abuse and violations of women's rights.
- To sensitize the public and raise awareness about women's rights.

Principal change/outcome objectives:
Awareness/knowledge
Advocacy/policy change

Key strategies:
Education and training

The project received $50,000 from the Trust Fund for the duration of one year, with 10% set aside for monitoring and evaluation costs.

3 Title: Researching stories of women living with HIV/AIDS from the 1994 genocide

In Rwanda, the everyday care of people living with HIV/AIDS largely falls on the family or the community. However, women who were raped during the 1994 genocide and are now living with HIV/AIDS lost most of their family and friends in the genocide leaving them isolated from the support and social network vital to their survival and well-being. Those with family often suffer deprivation and neglect due to stigmatization. Though there are many survivor groups in Rwanda, none deal specifically with the needs of these women.

Project objectives:
- To raise awareness on the link between conflict and post-conflict and VAW and the spread of HIV/AIDS in Africa.
- To raise awareness of the problems faces by survivors of rape living with HIV/AIDS.
- To encourage other rape victims in Rwanda and in the region to tell their story and deliver insights into the rehabilitation of women victims and the lack of women's support service in Rwanda and other countries.
- To reveal the enormous gaps in the provision of treatment of HIV/AIDS patients in Rwanda.
- To influence government and international policy and programs to show greater commitment to HIV/AIDS prevention and treatment.

Principal change/outcome objectives:
Awareness/knowledge
Advocacy / policy change

Key strategies:
Research and documentation
Raise public awareness

The project received $50,000 from the Trust Fund for the duration of one year, with 10% set aside for monitoring and evaluation costs.

4 Title: Degree of Commitment to Gender Equality in the Algerian Population

The project proposal has mainly investigative objectives, that is, it seeks to determine the extent of resistance to egalitarian gender relations by means of a survey. The purpose of the survey is to provide information that can better guide the strategies of NGOs involved in the fight for gender equality. Project activities include the finalization of a questionnaire applicable to three Maghreb countries (initially developed with an earlier Trust Fund grant), the launching of a survey in 10 regions on a sample of 1200 women, men and youth; discussion of the findings; and publication and dissemination of the results. The long term objectives are to contribute to a regional reform process that includes legal revisions. The direct beneficiaries of this initiative are the NGOs, associations and research institutions that work to eliminate violence against women. The final report and "impact evaluation" came in the form of a description of activities and a brief synopsis of the results which showed that women had a greater inclination toward egalitarian relations among women than among men.

Project objectives:

- To provide the knowledge base and develop tools to enable national NGOs, researchers and activists to adapt current strategies in advocating for gender equality and elimination of violence against children and women in Algeria.
- To launch a debate with communities and local NGOs in the 10 rural and urben regions targeted, by sharing and discussing the findings on the survey on the "Algerian Population's Degree of Commitment to Gender Equality".
- To issue a publication and disseminate findings to the public through media.
- To provide the research material for the publication of a comparative study on Gender Equality

Principal change/outcome objectives:
Awareness/knowledge
Advocacy/policy change

Key strategies:
Research and documentation
Educate and launch public discussion

The project received $48,800 from the Trust Fund for the duration of one year, with $1,023 set aside for monitoring and evaluation costs.

5 Title: Educational Campaign for the Prevention and Elimination of Violence Against Women

The overall objective of the project is to address the problem of violence against women in dating relationships. The project seeks to achieve the following specific objectives:

Project objectives:

- To determine the nature, extent and prevalence of abuse of young women by boyfriends in the high school population;
- To provide students with information and skills to protect themselves and/or prevent the occurrence of gender-based violence in schools;
- To provide information for teachers and parents regarding the incidence of violence and measures necessary to combat the problem;

- To equip teachers and parents with ways of communicating effectively with their students and children respectively;
- To influence government education policies by ensuring a safe environment for girls and young women;
- To contribute to a broader community initiative aimed at challenging gender discrimination and reducing violence against women and girls.

Principal change/outcome objectives:
Awareness/knowledge
Advocacy/policy change

Key strategies:
Research and documentation
Education through classroom discussion and conferencing
Training for teachers
The project received $62,000 from the Trust Fund for the duration of one year, with 10% set aside for monitoring and evaluation costs.

6 Title: Monitoring the 1994 Domestic Violence Act

In response to findings that the existing Domestic Violence Act in Malaysia has not resulted in significant improvements for survivors of domestic violence, the Women's Aid Organization has lobbied to revise the administration of the Act. In order to lend greater weight to its efforts, the organization proposed to conduct a monitoring and research study, documenting the treatment of victims of domestic violence in hospitals, police stations, and in courts. The objective of the study was to provide an advocacy tool that would be instrumental in bringing about change in the implementation of the law. The findings of the report were to be used as a basis for a workshop during which a strategy for improving current laws and procedures could be discussed with policy makers, various agents and decision makers. Target beneficiaries of the initiative are staff in offices that serve as points of contact for victims of domestic violence, women's groups and other non-governmental organizations that advocate on behalf of women, and ultimately, victims of domestic violence.

Project objectives:
- To produce a report which can be used as an advocacy tool to bring abut change in both the substance of the law and in administrative procedures in implementing the law.
- To identify loopholes in early detection of VAW and in legal instruments and procedures used to date.

Principal change/outcome objectives:
Awareness/knowledge
Advocacy/policy change

Key strategies:
Research and documentation
Public discussion

The project was awarded $42,880 over the period of 18 months, with 10% of the total budget allocated to evaluation costs.

7 Title: The Empowerment of Women in Tribal India: Capacity Building for Tribal Women Leaders to Respond to and Prevent Violence Against Women

The project challenges what it perceives as a widespread misconception that tribal women in India enjoy more egalitarian conditions and liberties. It attempts to address specifically the adverse effects of "development" processes on tribal women who are penalized by technological progress. In particular, they drive women out of productive activities, leaving them with decreased access to or control of natural resources and in increasing poverty. The diminishing economic role of women leads also to increased exploitation and violence against women within their communities. The project aims to build capacity and leadership among these women, sensitize tribal leaders to the issues, strengthen networking processes for advocacy purposes, and build public awareness through the dissemination of study and project results. Specifically, the project's strategy emphasizes the linking of micro and macro concerns of tribal women through a networking process at the local, regional and national levels.

Project objectives:
- Increased awareness of adivasi panchayat leaders on VAW and women's human rights issues
- Institutional change at local and district level
- Expansion of gramsabhas/panchayats powers to deal with issues relating to VAW
- Protection of tribal women against exploitation by non-tribal men
- Increase the capacity of tribal women and their leaders to defend themselves

Principal change/outcome objectives:
Awareness/knowledge
Advocacy / policy change
Empowerment

Key strategies:
Research and documentation
Training and capacity building

This two-year project received $50,000 from the Trust Fund, of which $4,600 were designated for impact assessment.

8 Title: A Field Study to Research and Document the Prevalence of Domestic Violence Against Women

The project proposal claims that there is no reliable data to back up the belief that VAW is widespread in Jordan. The objective of this project is to raise awareness of VAW through the publication of a field study that documents different forms of VAW in Jordan.

Project objectives:
- To identify the magnitude of the problem of violence against women in Jordan
- To identify the official procedural measures taken in dealing with cases of VW in Jordan
- To explore preventive and remedial measures and rehabilitation programs for helping victims of VAW
- To provide statistical data and information regarding VAW in Jordan
- To raise awareness of decision makers and society at large regarding the problem of violence against women and devise legislative and educational measures to prevent and eliminate it.

Principal change/outcome objectives:
Awareness/knowledge

Key strategies:
Research and documentation

The project was awarded $16,284 over the period of 1 year, with 10% of the total budget allocated to evaluation costs.

9 Title: Advocacy for Women's Human Right to a Life Without Violence

The proposal responds to a situation in which the incidence of domestic violence and other forms of VAW are on the rise in Bosnia, after a period of social and structural devastation. NGOs and institutions such as law enforcement, courts and medical providers are inadequately equipped to deal with this trend and need better training and information sharing if the trend is to be reversed. Based on a survey of attitudes and services available to victims of abuse, Infoteka, a women's advocacy group, proposes to conduct educational and training workshops for local governmental and non-governmental actors to promote needed policy revision and the development of stronger relationships among service providers.

Project objectives:
- To provide better legal, medical and psycho-social services to victims of violence while working to institute policy and legal changes which will help eliminate violence against women in Bosnia
- Increased awareness of attitudes and services available to victims of VAW
- Establishment of a women's lobby and information-sharing network

Principal change/outcome objectives:
Awareness/knowledge
Advocacy/policy change
Provision of services

Key strategies:
Research and documentation
Education and training
Building public awareness
Networking

This project received $48,900 from the Trust Fund for the period of one year, of which $513 were set aside for evaluation purposes.

10 Title: Women's Rights = Human Rights

Every year, significant numbers of Moldovan women arrange to work abroad where they often become victims of trafficking. Those that stay at home, on the other hand, suffer discrimination and are often subjected to domestic violence. Rural patriarchal traditions, stereotypes prevalent among the general public, insufficient information about women's rights, and a lack of legislation to support women, all contribute to women's vulnerability to domestic violence, trafficking, and other forms of violence against women in Moldova.

Project objectives:
- To educate members of NGOs and law enforcement officials about women's rights, with specific attention to trafficking and domestic violence.

- To increase cooperation between NGOs and law enforcement officials in protecting women's rights.
- To promote dialogue between different sets of actors – various social institutions, civilian and police/security personnel, men and women – for promotion of women's rights.
- To increase the exchange of information, experiences and methodologies with women's NGOs outside Moldova working on violence against women issues.

Principal change/outcome objectives:
Awareness/knowledge
Advocacy/policy change

Key strategies:
Education
Networking and information exchange

This project received $44,700 from the Trust Fund for the period of one year, none of which had been earmarked specifically for evaluation.

11 Title: Trafficking in the Commonwealth of Independent States (CIS) Countries

As follow-up to prior activities and its involvement in the drafting of new legislation on trafficking, Syostri, a regional NGO based in Russia, is "looking to strengthen networks and cooperation of groups in Russia, Ukraine, Armenia, Kyrgyzstan and Uzbekistan as a tool to raise awareness and prevent the development of trafficking and to increase regional NGO capacity to lobby appropriate institutions and mechanisms for policy and legal change". Syostri has also been focusing on the rise of new risk groups, including migrant women from the CIS and internal migration within Russia itself". The new initiative proposes the creation of an Informational and Training Center on Trafficking in Women in the CIS countries, the incorporation of CIS Women NGOs into a network and the facilitation of coordinated efforts in combating trafficking in women and children.

Project objectives:
- To eliminate trafficking in women and to render assistance to victims, through strengthening cooperation between CIS WNGOs and international organizations, governmental law enforcement and migration agencies, and mass media.
- To disseminate information, best practices and experience on compating trafficking in women and, in particular, on lobbying for modification of national legislation according to newly adopted international documents such as the UN Trafficking Protocol and Human Rights Standards on the Treatment of Trafficked Persons.

Principal change/outcome objectives:
Awareness / knowledge
Service provision
Advocacy / policy change

Key strategies:
Disseminate information
Training and capacity building

This project was awarded $75,750 for 16 months, of which $7500 were designated for impact evaluation.

12 Title: Partner Relations of Young Women with Special Emphasis on Issues of Relational Conflict

Project Summary:
The project addresses the problem of abuse or violence against young women who have little experience in managing relationships and who, influenced by traditional gender stereotypes, feel powerless to defend themselves.

The general objective of this project is "to contribute to the ideological changes in gender relational context and the creation of new more egalitarian male-female relationships in a social climate that rejects force and humiliation and encourages mutual respect, dignity and understanding".

The specific objective is a) to collect the data necessary to prepare an education and awareness raising program adjusted to the specific needs of the social, cultural and economic environment, and b) the implementation of the awareness-raising program itself. The awareness raising program will have the following four components: 1) workshops in schools and neighborhoods, 2) production of radio programs, 3) creation of a support network, 4) production and dissemination of materials on egalitarian partner relations in schools, workplaces, neighborhoods and youth groups..

Project objectives:
- To increase awareness, assertiveness and self-esteem of young women in relation to their partners.
- To educate women to make the connection between their individual problems and the broader cultural, social and economic environment.
- To develop communication and conflict resolution skills of young women in relationships, initiating and stimulating public discussion on male-female relations and the need for a structural change in gender relationships.

Principal change/outcome objectives:
Awareness/knowledge
Empowerment
Advocacy/policy change

Key strategies:
Research and documentation
Training and capacity building
Creation of support network

The project received $43,580 from the Trust Fund for a period of one year, with $4,000 set aside for impact assessment.

13 Title: Sexual Violence During Armed Conflict in Peru

Given the context of Peru's war that ended in 1996, the purpose of this project is to urge Peru's Truth and Reconciliation Commission (TRC), the national agency established to investigate human rights violations and formulate proposals for reparations of victims. The project's aim is to ensure that a gender perspective be incorporated into the work of the TRC and that female victims of rape and assault during the war be included in proposals for reparations, recovery and support.

Project objectives:
- To gather evidence on the use of rape as weapon during the Peruvian armed conflict
- To place violence against women on the agenda of the TRC, as well as civil society organizations
- To support the recovery of survivors of rape and sexual assault during the Peruvian armed conflict

Principal change/outcome objectives:
Awareness/knowledge
Advocacy

Key strategies:
Research and documentation
Training and capacity building
Lobbying
Education through seminar

The project received $50,000 from the Trust Fund for a period of one year, of which 10% were set aside for impact assessment.

14 Title: Changing Adolescent Attitudes on VAW

The project aims to address the problem of violence against women by working with secondary schools students in order to open lines of communication with their peers and with other males in the community.

Project objectives:
- To design implement and evaluate a program to train adolescent peer educators in a local secondary schools to give talks and/or workshops about violence against women to other students.

Principal change/outcome objectives:
Awareness/knowledge
Attitudes

Key strategies:
Awareness raising / education
Training and capacity building

The project received $43,927 from the Trust Fund for a period of one year, of which $5,400 were set aside for impact assessment.

15 Title: Program of Technical Assistance to Brazilian Mercocities Local Governments – Actions and Public Policies to Eliminate Violence Against Women

The Mercocities network is an organization that gathers mayors from large urban centers of Mercosur (the Southern Common Market) whose task is to promote the strengthening of local administrations as logical and natural counterpart to globalization. It is integrated by different cities of Argentina, Bolivia, Brazil, Chile, Paraguay and Uruguay. The general objectives of this project are to "support the formulation of integrated programs and plicies to eliminate violence

against women taking into account the globalized economic development and the consequent creation of the Latin-American countries markets.

The specific objectives are to do research, disseminate the findings, supply information on the approach and concepts related to violence against women and policies to eliminate VAW, foster interaction at government levels, provide support to existing Mercocity initiatives and assist the development of new initiatives.

Project objectives:

- To improve range and effectiveness of municipal-level initiatives on VAW by providing tools, strategies, and on-going technical assistance.
- To increase the access of Mercocities governments to national resources to combat VAW by fostering their interaction with the State Secretariat on Human Rights.
- To promote and facilitate the development of actions, programs, and policies to address violence against women in Mercocities currently lacking such initiatives.

Principal change/outcome objectives:
Service provision
Awareness/education
Advocacy

Key strategies:
Provide information based on research
Training and technical assistance

The project received $110,000 for a period of 22 months, and allocated $10,000 to impact evaluation.

2 UNIFEM Criteria for Evaluating TF Proposals

1. **Organizational Strength**. Applicants should demonstrate that they have both the experience and capacity to implement the project effectively, and manage funds. Preference is given to organizations based in the country where they will be working.
2. **Financial Viability**. Budgets should be feasible and should specify how the Trust Fund's contribution will be utilized in terms of activities. Funds should not be used to support ongoing organizational needs. Personnel should be no more than 40 per cent of the requested budget, while the administrative costs should not exceed 10 per cent. 10 per cent of the budget should be dedicated to impact assessment activities.
3. **Project Design**. A project should be cohesive with congruent relationships between the objectives, activities and budget. Activities should be discrete and achievable within stated timetables. The expected results of the project should be clearly stated and described in both quantitative and qualitative terms when possible, with clear steps outlined in the impact assessment for measuring corresponding indicators.
4. **Innovation.** Strategies and methodologies should strive to break new ground thereby contributing to a larger body of knowledge on ending violence against women. They should build on previous experience not only of the applicant, but also of other organizations working in related issues and strategies. Research projects should address emerging issues on VAW and its causes and consequences with clear plans for utilizing findings to stimulate change among target groups.
5. **Replicability**. There should be a potential for replication that can inform related initiatives, and thereby broaden the potential long-term impact of the project. Plans for sharing results and lessons of the project should be described.
6. **Partnerships & Linkages**. Projects should strive to forge synergistic partnerships between implementing organizations, institutions, and other groups, and bridge informational gaps between communities and governments. There should be confirmation of interest and availability of implementing partners and/or consultants and target audiences where feasible, especially governmental organizations.
7. **Sustainability**. Proposals should reflect substantial knowledge of the external environment, and address potential obstacles and openings. Strategies should be designed keeping in mind immediate results, but also should seek to put in place mechanisms that contribute to the continued involvement of target groups and/or partners beyond the life of the project. Institutions that can support the initiative in the long run should be identified early on. For example, training programs targeting law enforcement should involve professional schools or academies. Furthermore, strategies for sharing results and lessons of the project and expected follow-up activities should be articulated.