**Evaluation of International Humanitarian Operations**

by Marc Brenman

IDARE LLC

[mbrenman001@comcast.net](mailto:mbrenman001@comcast.net)

(US) 240-676-2436

Copyright 2015

Version of 9/9/2014

“The basis of true accountability, learning, and improvement is knowing in a timely fashion what people, processes, and program factors were effective in achieving the result.” (Chris Blattman; Impact Evaluation 2.0; 14 February 2008)

“I openly admit that some of the initiatives I have been involved in operationally did not work out as well as we wanted and in some cases fell far short of intentions for various reasons, ....in others we found ourselves inadvertent pawns in the politics and agendas of others.” (Chris Nixon, long term humanitarian worker, LinkedIn, 2/11/2013)

**Introduction**

“The purpose of humanitarian assistance is to enable people to survive crises and disasters, with an impartial focus on those who are most at risk…About 350 million people are affected by humanitarian crises every year, a significant proportion of whom are older people and people with disabilities, who have specific needs related to their age and disability.” (Jennifer Baird, Jessica Dinstl, Marcus Skinner and Jo Wells; A study of humanitarian financing for older people and people with disabilities, 2010–2011; HelpAge International, London and Handicap International, Lyon; 2012; [www.helpage.org](http://www.helpage.org))

Berkowitz and Cashman, in surveying forty Massachusetts-based healthy communities programs, found, “Almost all initiatives had engaged in some form of evaluation, but such evaluation tended to be irregular, partial and nonsystematic. There may have been good reasons: lack of time, lack of knowledge, and lack of qualiﬁed outside help. But, for most initiatives, evaluation was not a priority.”

(Berkowitz, W., and Cashman, S. “Building Healthy Communities: Lesson and Challenges.” Community, Fall–Winter 2000, 3 (2), 1–7)

**The Problem**

A report, "Proof of Concept – Learning from Nine Examples of Peacebuilding Evaluation," co-authored by USIP's Andy Blum, takes a look. Blum and his co-author, Melanie Kawano-Chiu, write in the introduction:

In a field that specializes in dialogue, consensus building and finding solutions to complex challenges, surprisingly absent is an open exchange on a fundamental element of peacebuilding: the determination of whether or not an intervention has ‘worked.'

**Evaluation**

Evaluation is the process of determining the value or amount of success in the achievement of predetermined objectives. This includes at least the following steps:

formulation of the objectives;

identification of the proper criteria to be used in measuring success;

determination and explanation of the degree of success; and

recommendations for future program activities.

Evaluation is the systematic and objective assessment of an on-going or completed project, program, or policy, including its design, implementation, and results. The aim is to determine the relevance and fulfillment of objectives, efficiency, effectiveness, impact, and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decisionmaking process of both recipients and donors. Even the most well-designed programs often need to be adjusted mid-stream. After an initiative is launched, emphasis shifts to regularly assessing its progress.

An organization should

gauge whether initiatives are on track to meet their stated goals;

identify necessary adjustments to strategy; and

ultimately assess return on investment.

Identify what works and what doesn’t in complex initiative design;

improve its ability to develop and manage program initiatives; and

make the organization more effective.

Historically, nonprofit organizations have been reluctant to invest in sophisticated evaluation systems since these systems have often been viewed as additional forms of overhead. Organizations can evaluate an individual program for its perceived effectiveness over time, making use of control groups and statistical analysis to measure certain outcomes.

The field is sometimes called “**Monitoring and Evaluation,” or “M+E.”**

**Evaluation is often not done at all, or not done well.** “[D]espite evaluation having been firmly on the peacebuilding agenda for a decade, the average peacebuilding team and agency lacks the knowledge and skills necessary to establish and manage quality evaluation processes…Underpinning this general lack of adherence to quality standards is a more fundamental issue: the distortion of the purpose of evaluation away from learning and accountability and towards fundraising, public relations and programme justification. The way this plays out in practice is a direct challenge to the espoused peacebuilding values of honesty and of the ends never justifying the means. (Cheyanne Scharbatke-Church; 'Evaluating peacebuilding'; [Berghof Handbook for Conflict Transformation](http://www.berghof-handbook.net/all); v. 2, 2011)

# Unfortunately, monitoring and evaluation often aren’t done well. For example, Erik Kijne, Senior moderator & trainer at PCM Group, has noted, “we live in the understanding that the 'monitor' or 'evaluator' should mainly facilitate the process and have the 'end-users' determine and agree on the indicators they consider important, participate in the collection of data…even have a crucial role in the analysis and interpretation of the findings.” (LinkedIn, [PCM Group, Project Coaches and Moderators](http://www.linkedin.com/groups?home=&gid=1578607&trk=anet_ug_hm), 2011)

“[I]n some cases perhaps a donor may not want an objective analysis and their ToRs are designed to support their own internal interests.” (Tim Hudson, LinkedIn, 2011)

**Why Evaluate?**

Evaluation is crucial to good humanitarian work: Formative research yields insights for program design; process research helps correct tactics that are off-target; and reliable outcome data tell us whether the effort produced desired changes. Evaluation is not only integral to effective humanitarian practice, it’s essential to growing a constituency for the work among social change practitioners and funders.

Assessment identifies policies, resources, and circumstances in the program’s context that may help or hinder implementation.

Evaluation results also provide the foundation for good stories, lending evidence of efficacy and cost-effectiveness. The power of story telling is a hot topic in the business literature. In Made to Stick, Chip and Dan Heath (2007) suggest that stories drive action through simulation (what to do) and inspiration (the motivation to do it). And some ‘‘springboard stories’’ help people see how an existing problem might change.

Evaluations contribute to transparency and accountability.

“Given the high turnover of staff in emergency situations, evaluations provide a means for organisations to retain and build institutional memory.” (Alistair Hallam; Evaluating Humanitarian Assistance Programmes in Complex Emergencies; Good Practice Review; Relief and Rehabilitation Network; September 1998; ODI) But only if the evaluation captures that memory, including how things were done, and lessons learned.

“In the face of the global economic downturn, policymakers and politicians no longer have the economic or political capital to fund projects without evidence of their efficacy. As a consequence, implementers and donor organizations alike are under increased pressure to demonstrate the utility of the programs they produce or fund.” (Amelia Arsenault, Sheldon Himelfarb, and Susan Abbott; Evaluating Media Interventions in Conflict Countries; US Institute of Peace; October 2011; http://www.usip.org/publications/evaluating-media-interventions-in-conflict-countries-0)

**The Need for Evaluation**

Humanitarian operations should be evaluated for feasibility, equity, effectiveness, efficiency, ethical and moral soundness (including lack of corruption), cost (including opportunity cost), sustainability, resiliency, scalability, and operating at cross-purposes.

**Measures drive improvement.** Teams of social service providers who review their performance measures are able to make adjustments in services, share successes, and probe for causes when progress comes up short — all on the road to improved outcomes.

**Types of measures** include

process measures to show whether an action was completed correctly;

outcome measures to show actual results;

experience measures to show how beneficiaries’ perspectives;

structural measures to reflect conditions under which services were provided; and

composite measures to combine the results of multiple measures to provide a comprehensive picture.

“Evaluations should collect both qualitative and quantitative data to develop a holistic picture. Qualitative data provide contextual information and help to explain the statistical data collected. Qualitative data can be collected through

interviews,

observations and

written documents;

quantitative data can be collected through

surveys and

questionnaires.

Evaluations should provide a comprehensive appraisal of

human,

material and

financial inputs.” (Minimum Standards for Education in Emergencies; Inter-Agency Network for Education in Emergencies; 2004; UNESCO; [www.ineesite.org](http://www.ineesite.org))

**Performance Accountability**

Performance accountability is defined as the mechanisms by which individuals or organizations are held accountable for meeting specified performance-related expectations. These are often categorized a quantity, quality, timeliness, cost, and customer service. They are also often divided by outputs and outcomes. **Outputs are units of delivery**, such as visits or widgets produced. **Outcomes are what has really happened or changed for the better** due to the efforts of the organization. These efforts have to be distinguished from demographic or random changes, or changes due to the work of other organizations. Part of accountability is that someone or some team bears the burden for lack of performance.

**Key criteria for developing performance measures** can be implemented, for example, including:

Develop a minimum set of performance measures that can be linked to a limited number of high-level goals and consistently applied across all the agencies, NPOs, and NGOs operating in an area.

**Minimum accountability requirements**. “The foundation’s proposal and reporting formats were specifically formulated to break down language and conceptual barriers that often serve to exclude community-based organizations from more typical funding mechanisms.  For example, rather than having to provide abstract objectives and a logframe, organizations were straightforwardly asked to respond to about ten questions such as

“What are you trying to achieve?” and

“What do you plan to do in this proposal to bring about these changes?”

At the reporting stage, the same occurred with such questions as

“What were you able to achieve?” and

“How do you know if you are making progress on your goals?”

Every attempt was made to de-technicalize development jargon and offer questions that were easy to translate in local languages, including financial reporting.” (Maryline Penedo; Small is Beautiful; no date; <http://www.how-matters.org/2011/01/13/small-grants-part-2/>; accessed 10/10/2011)

**Develop measures that demonstrate progress over time**, rather than measures tied to short-term targets.

Develop measures that emphasize incentives, training, and support, rather than penalties, as a preferred way to advance performance. However, **it is important not to reward poor performance.**

To measure performance, an organization, as part of its strategic plan, must develop goals, or desired outcomes, by which to determine program success.

**Pick the Right Proposal in the First Place**

**Did the project meet basic eligibility criteria?**

**Was it funded due to earmarks?**

“After a couple of decades experience, I came to the conclusion that proposal evaluation is something that one has to learn how to do. Important aspects are

assuring the neutrality of staff,

selecting good outside reviewers,

clearly stating the criteria for selection,

providing good guidelines for proposal preparation,

being sure that staff understand potential conflicts of interest of reviewers,

reviewing the review process with a willingness to withhold judgment of proposals that are inadequately reviewed,

providing good feedback to proponents of projects, and

dealing with the issue of confidentiality of reviewers.

It is hard to get to the right balance between assuring that proposals have sufficient information that they can be adequately reviewed and not demanding too much time and effort from the proponents in the preparation of the proposal.”

(Jennifer Lentfer; 1/14/2011; How Matters; <http://www.how-matters.org/2011/01/13/small-grants-part-2/>)

**Importance of Defining Benefits**

“The absence of formal cost-benefit analyses of outcome evaluations, including experiments, is glaring. Part of the problem lies in defining the benefit. Benefits are often not clear unless the evaluation is used, and "use" is often neither well documented nor well understood. An experiment's finding that a new program was unsuccessful, for instance, might imply an increased budget by way of salvaging a product in public demand, or a decreased budget in the interest of spending resources on more promising projects. Narrowly defined "use" may then be misleading and there is some risk of defining use so broadly as to make it meaningless. The difficulty of specifying political decisions beforehand exacerbates the problem.”

(Robert F. Boruch; Social Experiments : Education)

**Goal Setting**

Many organizations set goals for themselves. The goals should be based on realistic criteria, and not plucked from thin air. Organizations sometimes evaluate themselves against the goals they have set. Third party evaluators do this also. But often, questions about the appropriateness of the goal are not asked. The following should be evaluated:

Is the goal-setting realistic?

What was it based on?

How was it set?

Was it based on needs?

How were these needs identified?

Is it quantified?

Does it use ratio measures over time?

Is the goal vague or specific?

Is the goal just aspirational?

Is achievement of the goal dependent on factors outside the control of the organization?

How does the organization know if the goal has been achieved?

Are other organizations working on the same goal?

Was the particular goal picked primarily to obtain money from a funder?

**Are the interventions open to graceful failure?** That is to say, if the intervention is ended, will conditions on the ground be no worse off than they were before the intervention started? Or have the local people been deprived of the ability to make a living or support themselves? For example, were traditional forms of agriculture changed in favor of imported and expensive seeds?

**Are the perceptions of the local people considered?** “[T]here can often be considerable discrepancy between the agency’s perception of its performance and the perceptions of the affected population and beneficiaries. Experience shows that interviews with beneficiaries can be one of the richest sources of information in evaluations of humanitarian assistance. Interviews with a sample of the affected population should be a mandatory part of any humanitarian assistance evaluation.” (Alistair Hallam; Evaluating Humanitarian Assistance Programmes in Complex Emergencies; Good Practice Review; Relief and Rehabilitation Network; September 1998; ODI)

**Have the local people been protected?** “Evaluations of humanitarian assistance programmes should include an assessment of measures to provide protection to the affected population. In many emergency situations, the first needs of a population under threat may be protection from murder and harassment, as well as from discrimination that can lead to exclusion from basic services.” (Alistair Hallam; Evaluating Humanitarian Assistance Programmes in Complex Emergencies; Good Practice Review; Relief and Rehabilitation Network; September 1998; ODI)

“[T]he evaluation culture I envision is one that will emphasize fair, open, **ethical and democratic** processes…The data from all of our evaluations needs to be accessible to all interested groups allowing more extensive independent secondary analyses and opportunities for replication or refutation of original results.”

**Was a real problem addressed?** Or was the issue only a problem in the eyes of the intervenor? There should be no sacred cows among interventions. This is especially difficult when dealing with interventions by religious organizations. This is discussed in more detail below.

(William M.K. Trochim; Research Methods Knowledge Database; 10/20/2006; <http://www.socialresearchmethods.net/kb/evalcult.php>; accessed 6/12/2010)

**Have measures been taken to mitigate potential negative consequences of the program?**

**Humanitarian operations should be humane.** “In a 2002 article in Human Rights Quarterly, Barbara Harrell-Bond asked if humanitarian work can be made "humane." Rather than providing what recipients need, humanitarian groups largely do what their donors want accomplished, argues Harrell. Between the donors and recipients there may be a huge cultural gap that makes donor goals unhelpful or even inappropriate. Worse still, Harrell suggests that recipients learn a self-esteem-lowering role in order to maximize the aid they receive, setting up a symbiotic humanitarian-helpless relationship that may prolong the need for assistance. Indeed, we can look to some areas and suggest it may make the need permanent.   
  
Given this, humanitarian groups have an important (and unenviable) mission in translating recipient needs into donor language and culture…donors…goals must be aligned with the recipient’s needs. That means more than…calling attention to a tragedy and asking for aid. A more sophisticated, culturally literate presentation of specific needs is appropriate. Done well, this may help ward off the donor fatigue that plagues relief efforts after the initial pictures of hopelessness are moved off the screens in favor of new tragedies. Where displacement is the issue, rather than thinking about temporary shelter and water/food, health care, etc., from the very beginning the discussion perhaps should be about how to repatriate the displaced.” (Alan Howe, LinkedIn, USIP Alumni Group; 2/25/2012)

**When to Evaluate? [Timeframe for Evaluation]**

**“**Metrics that focus on process and that take into account stages of development are important for organizations to learn from their work as well as to both justify one foundation’s investment and to encourage others to jump in. Thus, funders should consider including evaluation capacity from the beginning, utilizing evaluation strategies that provide immediate feedback, and basing evaluation on a model that recognizes phases of development. (**Making Change: How Social Movements Work - and How to Support Them;** Manuel Pastor and Rhonda Ortiz; March 2009; http://college.usc.edu/geography/ESPE/perepub.html )

“Evaluations can take place prospectively (prior to the start), concurrently (periodically throughout the life of the project), and retrospectively (after the project is completed)…Partners could discuss both the objective of the project as well as the methods for accomplishing these objectives. Evaluation literature suggests that partners will develop more effective project activities if they begin by looking at the impacts they hope to achieve.

**Timeframe for Evaluation**

“The trend toward requiring increasingly elaborate project monitoring frameworks may need to be refined. Generally, ex ante analysis of the poverty reduction impact of projects is highly speculative. Moreover, transport and energy interventions contribute to poverty reduction over an extended period, perhaps as much as 20 years. Attempts to measure impacts over a shorter period are likely to confuse short-term effects with long-term impacts and produce misleading findings. It might be better to do ex ante poverty analysis of the sector rather than the project level, focusing on identifying and understanding the broad impact channels and critical situational factors. It would also be useful to initiate long-term monitoring studies to track the effects and impacts of a small sample of transport and energy projects over a period of 10 to 20 years.”

(Cynthia C. Cook, Tyrrell Duncan, Somchai Jitsuchon, Anil Sharma, Wu Guobao; Assessing the Impact of Transport and Energy Infrastructure on Poverty Reduction; Asian Development Bank; 2005; http://beta.adb.org/sites/default/files/pub/2005/assessing-transport-energy.pdf)

**Evaluation Plan**

• Another useful tool is an **evaluation plan**, which typically includes:

• Evaluation goals

• Partner assessment

• Purpose of program and program logic model

• Focus of evaluation (stakeholder needs, evaluation questions and design, resources

considerations)

• Methods for data collection and analysis

• Evaluation products and planned use

• Management of evaluation project (timeline, budget)”

Draft PEPH Evaluation Metrics Manual; Partnerships for Environmental Public Health (PEPH); National Institute of Environmental Health Sciences; National Institutes of Health; <http://www.niehs.nih.gov/research/supported/programs/peph/materials/index.cfm>

**Option: Evaluate before the end of the intervention**

**Types of Evaluation**

**“Formative Evaluation**: Used in the planning stages of a program to ensure the program is developed based on stakeholders needs and that programs are using effective and appropriate materials and procedures. Formative evaluation includes: needs assessments, evaluability assessment (analysis to determine if your program’s intended outcomes are able to be evaluated), program logic models, pre-testing program materials, and audience analysis.

**Process Evaluation**: It examines the procedures and tasks involved in providing a program. What services are actually being delivered and to whom?” Process evaluation includes such things as: tracking quantity and description of people who are reached by the program, tracking quantity and types of services provided, descriptions of how services are provided, descriptions of what actually occurs while providing services, and quality of services provided, implementation evaluation.

**Summative Evaluation**: Investigates the effects of the program, both intended and unintended. “Did the program make a difference?”(impact evaluation) “Did the program meet its stated goals and objectives?”(outcome evaluation). Outcome evaluation can assess both short term outcomes, immediate changes in individuals or participants (such as participation rates, awareness, knowledge, or behaviour) and long term outcomes (some- times referred to as impact evaluation) which look at the larger impacts of a program on a community. An outcome evaluation can also analyze the results in relation to the costs of the program (cost-benefit evaluations).” (Effective Health Communication; THCU Collection of Articles, March 2010)

**Real Time Evaluation (RTE) =** RTEis analysis of an operation undertaken early-on in order to provide immediate learning. The purpose is to enable management to make changes and alterations as necessary during an ongoing operation in order to ensure that the operation is meeting its objectives. An RTE is different from a post facto evaluation that is made to ensure accountability in terms of how money is spent or to register the ultimate impact of what has been done. Such evaluations are done as mid-term exercises or very soon after the completion of an operation. The RTE takes place during the course of implementation (EPAU recommends that it starts as early as possible).

• Like monitoring, it may aim to be iterative rather than one-off, hence the idea of on-going evaluation.

• The time-frame is short, with each exercise typically lasting days, rather than weeks.

• The methodology pays the usual attention to secondary sources of information, but is then interactive. Most RTEs are carried out through field visits combined with headquarters meetings, although some have been based purely on telephone interviews with field-based staff.

• RTEs use internal ‘consultants’ rather than, or perhaps alongside, externals/independents. The number of team members varies from one to many, but may include sectoral or other specialists, local staff or consultants.

• Restricted use is made of the DAC criteria for evaluation , with a greater emphasis on process.

• The emphasis is on immediate lesson-learning over impact evaluation or accountability.

• ‘Quick and dirty’ results enable a programme to be changed in mid-course. This brings RTE closer to monitoring, which primarily tracks progress, than to evaluation, which makes value judgements.

# Bernard Broughton suggests looking at ‘the operation’s relevance and design, progress in achieving the operation’s objectives (i.e. results), any gaps or unintended impact, the effectiveness and efficiency of the mode of implementation, and the appropriateness and application of operational guidelines and policies’. (Maurice Herson and John Mitchell, ALNAP; Real-Time Evaluation: where does its value lie? Humanitarian Practice Network; December 2005; http://www.odihpn.org/report.asp?id=2772)

# Option: Self-Evaluation

# An organization can evaluate itself. For example,

**Voluntary self-testing** = “In 1996, Congress created a legal privilege for lenders for data gathered in voluntary self-tests to assess compliance with the Fair Housing Act and Equal Credit Opportunity Act. The results of voluntary self-testing do not need to be disclosed provided that appropriate corrective action is taken to address possible violations that may be discovered. This privilege was meant to create an incentive to self-test. 12 C.F.R. § 202 (2011).” (Jorge Andres Soto and Deidre Swesnik; The Promise of the Fair Housing Act and the Role of Fair Housing Organizations; January 2012; American Constitution Society for Law and Policy Issue Brief; https://www.acslaw.org/sites/default/files/Soto\_and\_Swesnik\_-\_Promise\_of\_the\_Fair\_Housing\_Act.pdf) See also self-assessment.

An International Monetary Fund Report said, “voluntary compliance and self-assessment are the foundation of modern tax administrations.” Carlos Silvani & Katherine Baer, “Designing a Tax Administration Reform Strategy: Experiences and Guidelines” (Washington D.C.: 1997) IMF Working Paper WP/97/30 at para. 30. [97 TNI 149-31 at para. 30.]

# Cost of Evaluations

One consultancy company that has carried out a number of evaluations recommends, as a rule of thumb, that 2% of total program costs should be assigned for evaluation (with 4% for monitoring). (Phil O’Keefe, ETC UK) This is supported by the Millennium Challenge Corporation, which has stated, “MCC has found that establishing high expectations and budgeting appropriately – often in the range of 2-4 percent of the total program budget – creates an environment within which independent evaluations of impact can be conducted as part of the core implementation plan.” (Frank Wiebe, Chief Economist, Millennium Challenge Corporation)

“The costs of performance monitoring are modest relative to impact evaluations, but still vary widely depending on the data used. Most performance indicator data come from records maintained by service providers. The added expense involves regularly collecting and analyzing these records, as well as preparing and disseminating reports to those concerned.”

(Evaluation Strategies for Human Services Programs: A Guide for Policymakers and Providers; Adele Harrell *with* Martha Burt, Harry Hatry; Shelli Rossman; Jeffrey Roth, and William Sabol; The Urban Institute; Washington, D.C; <http://www.ojp.usdoj.gov/BJA/evaluation/guide/documents/evaluation_strategies.html>)

The Coalition for Evidence-Based Policy has found examples of five well-conducted, low-cost RCTs, carried out in real-world community settings. Study costs range from $50,000 to $300,000, with random assignment itself comprising only a small portion of this cost (between $0 and $20,000). The studies all produced valid evidence that is of policy and practical importance. See Resource Chapter for contact information.

**What isn’t evaluation?**

**Monitoring.** “Monitoring systems should meet the on-going information needs of programme managers, and should not be seen as existing to provide information solely for evaluation purposes.” (Alistair Hallam; Evaluating Humanitarian Assistance Programmes in Complex Emergencies; Good Practice Review; Relief and Rehabilitation Network; September 1998; ODI)

“Evaluation is often distinguished from monitoring in terms of its purpose, its empirical base, its depth of analysis and its duration and frequency. Evaluation usually feeds more into policy making and strategic planning (rather than operational management decision making processes), involves in-depth data collection and analysis, and is consequently undertaken only a few times during the project life (such as at, or some time after, the completion of a project or programme) rather than on an periodically basis.” (European Commission, EuropeAid Co-operation Office; Result Oriented Monitoring Handbook; July 2009; http://ec.europa.eu/europeaid/how/ensure-aid-effectiveness/documents/rom\_handbook2009\_en.pdf)

Monitoring is a continuous function that uses the systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds, time, energy, and other resources. A process that involves collecting, storing, analyzing and evaluating information, to measure performance, progress or change.Monitoring will help to assess whether a program is being implemented as was planned. A program monitoring system enables continuous feedback on the status of program implementation, identifying specific problems as they arise.  **(**World Bank; Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners; Judy L. Baker; 2000; [http://siteresources.worldbank.org/INTISPMA/Resources/handbook.pdf](http://www.linkedin.com/e/3oya7g-griw0aso-3f/plh/http%3A%2F%2Fsiteresources.worldbank.org%2FINTISPMA%2FResources%2Fhandbook.pdf/5MFy/" \t "_blank))Monitoring race equality involves collecting, storing, analyzing and evaluating information about the racial groups to which people say they belong. Monitoring is also applied to collecting and analyzing information about people’s gender, disability status, sexual orientation, religion or belief, or age to see whether all groups are fairly represented.

“The term ‘**internal monitoring’** is often used to refer to monitoring that is undertaken by those responsible for project implementation. There is little to no organisational distance between the person in charge of data collection and analysis and the user of the monitoring results. Done with varying degrees of methodological elaboration, it gives the manager instant feedback necessary for the day-to-day operation. Internal monitoring and reporting often overlap.

**External Monitoring** on the other hand separates clearly the management and the monitoring function. Monitoring is done by organisationally independent monitors who present their final results to management.” (European Commission, EuropeAid Co-operation Office; Result Oriented Monitoring Handbook; July 2009; http://ec.europa.eu/europeaid/how/ensure-aid-effectiveness/documents/rom\_handbook2009\_en.pdf)

**Monitoring and reporting procedures =** Procedures not generally resulting in legally-enforceable remedies; monitoring and reporting procedures resemble "**audits" of government behavior** which results in non-binding recommendations. In some cases, the reporting resembles a self-inspection; governments report on their own compliance with human rights obligations or a monitoring body initiates the report on government behavior. We can think of monitoring in terms of increasing levels of involvement:

At the lowest level, ***observation****,* we have pure passive watching and inspection. At this level, monitors lack the mandate to judge the actions of the parties being monitored; they simply observe what is going on.

As involvement increases, monitors will begin to judge, to verify compliance with the treaty. In this situation, parties may not only observe actions in relation to the agreement, but also judge and report violations. This is ***verification***.

The last level is ***enforcement***. At this level, monitoring of an agreement involves not merely observing and compliance with the treaty, but also enforcing the terms of the agreement through positive and negative incentives.

**Auditing.**  Auditing is usually financial, and concerned with adherence to industry standard rules. “Audit can be distinguished from monitoring and evaluation by its financial, and financial management, focus. It is primarily an assessment of the legality and regularity of project expenditure and income and whether project funds have been used efficiently and economically and for the intended purposes.” (European Commission, EuropeAid Co-operation Office; Result Oriented Monitoring Handbook; July 2009; http://ec.europa.eu/europeaid/how/ensure-aid-effectiveness/documents/rom\_handbook2009\_en.pdf)

**Process**

**“Process evaluations** are geared to fully understanding how a program works and how it produces immediate results. It is an examination of procedures and tasks involved in implementing a program, and can also include an analysis of administrative and organizational aspects. Short term outputs are often evaluated in a process evaluation. These evaluations can be useful for accurately portraying to outside parties how a program operates so that it could be replicated elsewhere. Sample process evaluation questions include:

• Did the project target the right audience?

• Did activities progress as expected?

• Was a committee created as planned?”

Draft PEPH Evaluation Metrics Manual; Partnerships for Environmental Public Health (PEPH); National Institute of Environmental Health Sciences; National Institutes of Health; <http://www.niehs.nih.gov/research/supported/programs/peph/materials/index.cfm>

In the course of evaluating the success or failure of a program, a great deal can be learned about how and why a program works or does not work. The analysis of process can have both administrative and scientific significance particularly when the evaluation indicates that a program is not working as expected. The **analysis of process** should be made according to four main dimensions, beginning with:

1) the attributes of the program itself;

2) the population exposed to the program;

3) the situation or context within which the program takes place; and

4) the different kinds of effects produced by the program.

**Outcome Evaluations**

“**Outcome or impact evaluations** ask whether the program activities are bringing about the desired effects. Outcomes and impacts should not be confused with outputs. Outputs are the direct products of an activity, such as training sessions, brochures, websites, or videos. Outcomes and impacts are the benefits or changes resulting from the activities and outputs, such as reduced exposure to environmental contamination, changes in behavior or policy, or, ultimately, improvement in public health. As a general rule, logic models lay out short term outcomes that result within 1–3 years, long-term outcomes within 4–6 years, and impacts within 7–10 years. The timing varies by project. What is most important here is that the ultimate goal for partnerships is to produce outcomes and impacts that lead to improvements in health through a reduction in environmental health hazards.

**Impact evaluations** focus on questions of causality. (Evaluation Strategies for Human Services Programs: A Guide for Policymakers and Providers; Adele Harrell *with* Martha Burt, Harry Hatry; Shelli Rossman; Jeffrey Roth, and William Sabol; The Urban Institute; Washington, D.C; <http://www.ojp.usdoj.gov/BJA/evaluation/guide/documents/evaluation_strategies.html>)

**Why is there so little evidence of impact?** “Doing impact evaluation well is not easy or cheap. Studies that take impact, causality and attribution seriously take a long time to do and attract substantial costs – even more so in difficult contexts.

**Fund programmes, not studies**. In conflict-affected situations, donors are faced with a huge number of urgent humanitarian and recovery needs. Funding research may not be at the top of their list of priorities when there are other, more pressing things to invest in.

Assumptions of effectiveness can prove remarkably resilient. To many, it may seem obvious that giving people jobs in war zones is a good thing to do – why spend money on research that will simply tell us what we already know? Deductive logic such as this is certainly compelling, and often convincing, but research can turn conventional wisdom on its head.

The truth might hurt. If a donor has been funding programme x for several years, it may not be in their interest to then fund research that tells them they’ve been doing it wrong.

We are measuring impact! Many studies we came across in our review used ‘impact’ to refer to how well a programme functions in terms of its own design – i.e. was it completed on time? Was the right amount of, say, seed distributed? This may be one way of measuring success, but it doesn’t tell us anything about what the programme did for beneficiaries.”

(Richard Mallett; Mind the gap: the fragile state of the impact evidence base; Secure Livlihoods Research Consortium; 12/3/2012; http://securelivelihoods.org/blogpost/32/Mind-the-gap-the-fragile-state-of-the-impact-evidence-base)

**When conducting an outcome or impact evaluation**, partners might want to ask:

• Did the program achieve the desired outcome?

• What can others learn from the program?

For evaluation projects, the overall process generally involves an assessment of inputs, activities, outputs, outcomes, and impacts. These components are interrelated. Available inputs for a given project can influence the choice of activities, and likewise the activities can define the outputs.”

Draft PEPH Evaluation Metrics Manual; Partnerships for Environmental Public Health (PEPH); National Institute of Environmental Health Sciences; National Institutes of Health; <http://www.niehs.nih.gov/research/supported/programs/peph/materials/index.cfm>

**Program Evaluation**

**Program** **evaluation** is an assessment, through objective measurement and systematic analysis, of the manner and extent to which programs achieve intended objectives. The purpose of this program evaluation plan is to improve the analytic content of evaluations.

**Cost. [Cost of programs; cost evaluation; cost measurement; cost studies]**

**“Cost evaluations** address how much the program or program components cost, preferably in relation to alternative uses of the same resources and to the benefits being produced by the program.” (Evaluation Strategies for Human Services Programs: A Guide for Policymakers and Providers; Adele Harrell *with* Martha Burt, Harry Hatry; Shelli Rossman; Jeffrey Roth, and William Sabol; The Urban Institute; Washington, D.C; <http://www.ojp.usdoj.gov/BJA/evaluation/guide/documents/evaluation_strategies.html>)

“Cost studies are used to assess investments in programs by collecting information on:

1) direct program expenditures;

2) the costs of staff and resources provided by other agencies or diverted from other uses;

3) costs for purchased services; and

4) the value of donated time and materials.

Costs for the first two items usually include expenditures for staff salaries; fringe benefits; special training costs (if any); travel; facilities; and supplies and equipment that have to be purchased. The value of donated resources, which can be substantial, generally has to be estimated and requires careful documentation of the donation.” (ibid)

**Cost-Benefit and Cost-Effectiveness Analyses** compare a program’s outputs or outcomes with the costs to produce them. One method of cost-effectiveness, that used by MIT’s Poverty Action Lab, “shows the impact against a specific policy goal that can be achieved for a given expenditure (e.g.  additional years of education per $100 spent). All the impact estimates are based on the evidence from rigorous randomized evaluations. A number of judgment calls have to be made about how to measure costs and benefits. In most cases there is no single “right” answer—different approaches are appropriate in different contexts (for example which exchange rate or discount rate to use will depend on whether it is an international donor or developing country government that is making the decision about what to invest in).” (<http://www.povertyactionlab.org/policy-lessons/cost-effectiveness>)

This type of analysis conforms with program evaluation when applied systematically to existing programs and when measurable outputs and outcomes are monetized. Program evaluations are retrospective, quantitative assessments of existing programs. The purpose of evaluation is not merely to comply with funding guidelines or to lend an air of legitimacy to a research and development effort. Rather, the goal is improvement. Program evaluation differs fundamentally from research in the purpose of data collection and standards for judging quality. Research aims to produce knowledge and truth.  Useful evaluation supports action. A successful evaluation (one that is useful, practical, ethical, and accurate) emerges from the special characteristics and conditions of a particular situation — a mixture of people, politics, history, context, resources, constraints, values, needs, interests, and chance.

Cost per client.

Cost per unit of service.

“**Monetization of benefits** to individualsrequires assumptions about three matters, all of which are frequently controversial. First, the dollar value of the benefit may depend on personal values. Second, a dollar of benefit today is worth more than a dollar benefit realized next year. Thus, the benefits need to be time discounted, but by how much is a difficult question. Third, the beneficiaries need to be identified. Societal values become important when the beneficiaries differ in standing and perceived merit.”

(Evaluation Strategies for Human Services Programs: A Guide for Policymakers and Providers; Adele Harrell *with* Martha Burt, Harry Hatry; Shelli Rossman; Jeffrey Roth, and William Sabol; The Urban Institute; Washington, D.C; http://www.ojp.usdoj.gov/BJA/evaluation/guide/documents/evaluation\_strategies.html)

**Social costs averted**. “These are the savings to the public that result from avoiding negative outcomes…These estimates are difficult to derive and are often based on tenuous assumptions. To compensate for problems in the reliability of estimates, cost-benefit calculations normally use a range of benefits to place an upper and a lower bound on the probable returns to investments in the program**.** A more significant problem is that monetary values based on public costs for the negative outcomes among the general population may be poor estimates of the value of benefits among the program's *target* population…Other public benefits reflecting gains, not costs averted, are widely acknowledged, but rarely find their way into cost-benefit studies because there is no public consensus on their importance. Examples include improvements in the quality of life or the environment.” (ibid.)

**Efficiency.** Efficiency is a measure of how well economical resources/inputs (funds, expertise, time, etc.) are converted to results or outputs. It is concerned with producing the greatest total social value (as determined subjectively by individuals and as measured by economists either in markets or by using non-market methods) for the least possible social cost. It is achieved when all resources are in their most productive use (production efficiency), no mutually beneficial trades of goods and services are possible (exchange efficiency), and net value is maximized. No reallocation of resources or consumption can result in an increase in total value net of cost. (NOAA)

Efficiency is decreased by waste and **corruption**. Corruption is the abuse of entrusted power for private gain. There are several ways of looking at corruption. One way is to differentiate between "according to rule" corruption and "against the rule" corruption. Facilitation payments, where a bribe is paid to receive preferential treatment for something that the bribe receiver is required to do by law, constitute the former. The latter, on the other hand, is a bribe paid to obtain services the bribe receiver is prohibited from providing. (Transparency International) The abuse of public office for private gain is a commonly accepted definition of corruption. See, e.g., Nikolay A. Ouzounov, Facing the Challenge: Corruption, State Capture and the Role of Multinational Business, 37 J. MARSHALL L. REV. 1181, 1186 (2004). Corruption often takes the form of bribes demanded or hinted, and paid. It should be noted, however, that in some parts of the world and in some cultures, it is essential to participate in the corruption practices in order to get anything important or necessary done. An NPO or NGO can experience conflicts between the necessary and essential, and its own moral and ethical precepts. National law of the home country can govern also, as in the US’s Foreign Corrupt Practices Act.

Corruption, and beliefs about corruption in the home country of the NPO or NGO, can influence feelings about the effort and support for it. For example, Kaiser Family Foundation President Drew Altman has stated, “One of the strongest predictors of support for global health spending was the belief that aid would make a difference.  This means that documenting the impact of assistance and then communicating that to opinion leaders and the public is absolutely critical for advocates of foreign aid and global health.  We found that the public believes that almost half of every dollar we spend to help other countries is lost through corruption, so this is a formidable challenge.” While this figure is obviously exaggerated, organizations like Transparency International do provide estimates of corruption infrastructure in many countries, and rank them.

**“Connoisseurial” evaluation.** Connoisseurial evaluators are free to make use of objective and subjective measures. The connoisseur’s judgment is not used to supplant empirically based evaluation, but rather to supplant, amplify, and explain more traditional evaluation research (Bridglall, B. L. and E.W. Gordon. 2004. Creating Excellence and Increasing Ethnic Minority Leadership in Science, Engineering, Mathematics, and Other Technical Disciplines: A Study of the Meyerhoff Scholars Program at the University of Maryland, Baltimore County. New York: Institute for Urban and Minority Education, Teachers College, Columbia University; 16).

**Murder Boards and After-action Reports**

Used in the US military after a mission to help determine what went right and wrong. Staff are free to be frank.

**The four basic criteria for evaluation** are as follows:

* *Utility.* Utility involves taking steps to ensure that “an evaluation will serve the practical information needs of intended users.” As committee chair Daniel Stufflebeam observed, “an evaluation should not be done at all if there is no prospect for its being useful to some audience.”
* *Feasibility* — To be judged as feasible, an evaluation must be “realistic, prudent, diplomatic, and frugal.” Stufflebeam suggests an evaluation should not be conducted “if it is not feasible to conduct it in political terms, or practicality terms, or cost-effectiveness terms.”
* *Propriety* — Propriety involves ensuring that “an evaluation will be conducted legally, ethically, and with due regard for the welfare” of both the individuals participating in the evaluation and those who are affected by its results.
* *Accuracy* — To be accurate, an evaluation must “reveal and convey technically adequate information about the features that determine worth or merit” of what’s being evaluated. This criterion addresses the methodological appropriateness of both the evaluation plan and its implementation

**Performance**

**The performance or effect criterion measures the results of the effort rather than the effort itself.** This requires a clear statement of the project’s objective. How much is accomplished relative to an immediate goal? Did any change occur? Performance can be measured at several levels- for example, in public health, the number of cases found, the number hospitalized, the number cured or rehabilitated. Performance standards often involve several assumptions; however, in general, evaluation of performance involves fewer assumptions than evaluation of efforts.

**Efficacy**

A positive answer to the question, "Does it work?" often gives rise to the following question: Is there any better way to obtain the same results? Efficacy is concerned with the evaluation of alternative paths or methods in terms of cost and money, time, personnel, and public convenience. In a sense, it represents the ratio between effort and performance, output divided by input.

**Effectiveness**

Effectiveness is the extent to which a project’s objectives were achieved, or are expected to be achieved, taking into account their relative importance. It is used as an aggregate measure of (or judgment about) the merit or worth of an activity, i.e., the extent to which a project has attained, or is expected to attain, its major relevant objectives efficiently in a sustainable way and with a positive impact. A related term is efficacy.

“Many studies…employ weak, or at least limited, measures of effectiveness. For example, several are based on perceived effectiveness rather than observed effects…Self-reported effectiveness is problematic because it does not correlate well with actual behavior (Austin, Pinkleton, & Fujioka, 1999).” (Gerard Hastings, Martine Stead, and John Webb; Fear Appeals in Social Marketing: Strategic and Ethical Reasons for Concern; Psychology & Marketing, Vol. 21(11): 961–986 (November 2004);

**Connectedness**

This refers to the need “to assure that activities of a short-term emergency nature are carried out in a context which takes longer-term and interconnected problems into account.” (Minear, L. (1994), *The International Relief System: A Critical Review*, available from Thomas J. Watson Jr. Institute for International Studies (http://

[www.brown.edu](http://www.brown.edu)).

**Coverage**

This refers to the need “to reach major population groups facing life-threatening suffering wherever they are, providing them with assistance and protection proportionate to their need and devoid of extraneous political agendas” (Minear, L. (1994), *The International Relief System: A Critical Review*, available from Thomas J. Watson Jr. Institute for International Studies (http://[www.brown.edu](http://www.brown.edu)). “Coverage alerts evaluators that complex emergencies and associated humanitarian programmes can have significantly differing impacts on different population sub-groups, whether these are defined in terms of ethnicity, gender, socio-economic status, occupation, location (urban/rural or inside/outside of a country affected by conflict) or family circumstance (e.g. single mother, orphan). Programmes need to be assessed both in terms of which groups are included in a programme, and the differential impact on those included.” (Alistair Hallam; Evaluating Humanitarian Assistance Programmes in Complex Emergencies; Good Practice Review; Relief and Rehabilitation Network; September 1998; ODI)

**Coherence**

“This refers to the need for policy coherence, the need to assess political, diplomatic, economic and military policies as well as humanitarian policies, to ensure that there is consistency and, in particular, that all policies take into account humanitarian considerations.” (Alistair Hallam; Evaluating Humanitarian Assistance Programmes in Complex Emergencies; Good Practice Review; Relief and Rehabilitation Network; September 1998; ODI)

**Tools for Evaluation**

**Use of Social Media**

Social media are increasingly being used for evaluating humanitarian operations, as more people have access to cellphones and integrated devices. For example, Global Pulse (www.unglobalpulse.org) is an innovation initiative of the UN Secretary-General, harnessing today's new world of digital data and real-time analytics to gain a better understanding of changes in human well-being. ActionAid International is interested in this area, and is interested in collaborating. They want to look at Web 3.0 for monitoring their own programmes, generating feedback with people they work with via cellphones that can be analysed for patterns and trends. The American Evaluation Association is also interested. The Humanitarian Innovation Fund is using a suite of technologies – mobile telephony, internet, on-line communities and social media, geo-mapping – to foster beneficiary participation in development and humanitarian interventions, by encouraging beneficiaries to “express their demands, aspirations, engaging in the process of formulation of humanitarian interventions, planning, monitoring and evaluation”. (Kim Scriven)

It is important not to confuse the tool, social media, with actual delivery of services in improving the lives of real people. Delivering information is not the same as delivering needed aid. However, it is also true that some people may prefer the information and communication, and then use these tools to help themselves. Downsides include use of social media for propaganda, national governments seizing the means of communication, and social media being used for perverse purposes, like radio was used in Rwanda to lead people from one ethnic group to kill members of another. A tool is not inherently good or bad, and can be used for good, bad, and indifferent purposes.

**Product of the Evaluation**

**Evaluation Report =** Anevaluation report is a product of an evaluation. It should clearly identify the purpose of the evaluation, what was evaluated, how the evaluation was conducted, the data considered, the conclusions drawn and recommendations made and lessons identified. The report should explain how each conclusion (judgment on findings) derives from the findings, and their limitations. Recommendations should be linked to conclusions, and addressed to the appropriate hierarchical level (operational, strategic, etc.) for follow-up. As to the lessons, the conditions under which the knowledge can be transferred should be spelled out.

The report should include:

* an executive summary of conclusions and recommendations and lessons learned;
* information on the independent evaluation and the methodologies used;
* information on the selected contexts and on the history of the program implementation in the selected contexts;
* field observations and findings;
* prospects and options moving forward;
* Conclusions and Recommendations;
* lessons learned; and
* annexes (Terms of Reference; inception report; Work Plan; field visits; list of interviewees; eferences / documentation).

In addition to the final text, the report should contain:

* a title page;
* who did the evaluation;
* why it was done;
* the evaluator's background;
* the complete table of contents including annexes;
* a list of abbreviations used;
* a glossary if necessary.

**Code of Ethics for Evaluation and Evaluators**

The evaluator should not have been involved in the design and implementation of the program, nor have other conflicts of interest. See Appendix \_\_, Model Code of Ethics.

**What should not be included positively in an evaluation?**

The tax exempt status of the organization. Religious organizations have tax exempt status, for example.

Whether the organization has movie stars, radio or TV talk show hosts, or other celebrities in its leadership or support. Whether a very rich person provides support is a tough issue. For example, if the rich person helps fund the organization, that is an important source of revenue.

The flashiness or cuteness of the organization’s designs, mascots, or symbols. For example, some animal welfare organizations organize around cute “megafauna”—fluffy cute animals like polar bears or pandas.

**Foundational Questions**

How long has the organization been involved in this particular country/area?

Does the organization have adequate competence to deal with the problems presented?

Was it aware of the problems presented?

Was assistance provided on the basis of the agencies’ interests and resource situation rather than actual needs?

Did the organization frame the problem correctly? For example, it is consistently seen in the response to drought and famine disasters that these are perceived and dealt with as a question of food shortage rather than livelihood or infrastructure or building crises. This results in focus on the life-saving distribution of food and medicine, but limited focus on rehabilitation and the development of livelihoods or infrastructure to reduce the population’s vulnerability to new disasters (Steering Committee 2004). The ALNAP reports are skeptical of this type of “food aid” because it can create dependency rather than promote vulnerability-reducing livelihood and infrastructure development (ALNAP 2003). (See also discussion of resiliency.)

What have been the long term effects?

Does the organization have an exit strategy?

Have local people become dependent on humanitarian aid?

Did the organization made a strategic review at the crucial point when turning from relief to recovery?

**Qualitative research activities** are needed throughout the life of equity programs. This need begins with the planning stages through monitoring and evaluation.  Qualitative research methods can be used for the following planning and management tasks:

* identify and understand the beneficiary population’s overall priorities for action and the ranking of different sector issues (e.g., health, water, income, food, crop production) among priorities;
* identify and understand the beneficiary population’s specific priorities within a specific sector such as health;
* identify and understand the underlying reasons for problems before developing solutions;
* identify and understand the beneficiary population’s language, concepts and beliefs surrounding specific behaviors/situations targeted for change; and,
* assess stakeholder reactions to our programs to adapt implementation and evaluate (subjectively) the immediate effects of our program.

**Performance monitoring.** “provides information on key aspects of how a system or program is operating and the extent to which specified program objectives are being attained (e.g., numbers of youth served compared to target goals, reductions in school dropouts compared to target goals). Results are used by service providers, funders, and policymakers to assess the program's performance and accomplishments.: **(**Evaluation Strategies for Human Services Programs: A Guide for Policymakers and Providers; Adele Harrell *with* Martha Burt, Harry Hatry; Shelli Rossman; Jeffrey Roth, and William Sabol; The Urban Institute; Washington, D.C; http://www.ojp.usdoj.gov/BJA/evaluation/guide/documents/evaluation\_strategies.html)

**Has the organization’s work addressed real needs on the ground?**

Have local people been consulted as to what they see as their needs?

**Has local capacity been grown?**

Have local people been trained in the skills necessary to carry on the work of the organization?

Have local people been prepared to deal with future problems on their own?

Has there been handoff to local partners?

**Are relations with local partners optimized?**

**Are the local efforts sustainable?**

For example, when hospitals or housing are built in earthquake-prone areas, are they built according to seismic code? Are structures built in floodplains? Are they resilient?

Have structures been built in an accessible way to people with disabilities?

Are solutions capable of being maintained by local people with local resources?

Has the organization devoted resources to ensuring that the problem addressed will not arise again?

Have the local people and infrastructure become more resilient as a result of the organization’s work?

Has the community become mobilized in a positive direction?

**Did the organization take advantage of the opportunity to produce positive social change through advocacy?**

**Has the organization coordinated with other humanitarian organizations?** With the UN? With other international efforts?

Have joint evaluations been conducted, to encourage and promote constructive peer review, improve cost effectiveness and better capture attribution?

**What is the unique contribution of the organization?**

**How have the efforts of the organization contributed to long-term development?**

To self-reliance?

**Is institutional learning carried over from previous projects?**

**Are outcomes monitored?**

**Is there analysis of the social, cultural and institutional factors of success of projects**, or does the organization just do what it does?

**Have women’s rights and control of their bodies been supported?**

**Have minority rights been supported?**

Have services been provided irrespective of race, ethnicity, and religion?

**Have the rights of people with disabilities been supported?**

Are programs accessible to people with disabilities?

**Are there disadvantages, downsides, or negative externalities to the program?**

“Development initiatives in themselves are not risk neutral - they can increase hazards and exacerbate vulnerabilities if they are poorly designed or implemented, or if they cause serious environmental harm. Economists typically view a construction boom as a sure sign of a country's progress and growth. But sometimes newly constructed infrastructure simply invites disaster. Contractors often cut corners and do not execute engineering designs correctly…Poor adherence to building standards plus the pervasive corruption across South Asia also result in high levels of urban risk…Donor-led humanitarian responses sometimes sideline local leaderships and reduce the capabilities required for long-term resilience.” (Oxfam International; Rethinking Disasters: Why death and destruction is not nature’s fault but human failure; 2008; <http://www.preventionweb.net/files/1764_oxfamindiarethinkingdisasters.pdf>

South Asia Regional Centre, Oxfam (India) Trust)

This was true in Haiti, where many of the hospitals, medical clinics, and schools were not build to seismic safety standards.

Evaluating corruption is difficult, especially in gift-giving cultures, where little happens without gifts being presented.

**Impact Assessment**

**“**The fundamental tenet of impact assessment is the need to **compare the observed situation with the intervention to *what would have been* had there been no intervention at all**, i.e., the counterfactual. The difference in resulting outcomes between these two states constitutes the impact of the intervention. While the counterfactual cannot be observed or known with complete certainty, the concept of comparing observed outcomes to this hypothetical state underlies all valid approaches to assessing impacts. Valid comparisons imply that the net effect of interventions is isolated from all other extraneous or confounding factors that influence defined outcomes. …the major challenge in impact assessments is to estimate the effect of programs after netting out extraneous factors that affect outcomes. These factors may include specific events or long-term trends in industries, regions or countries.”

**Methods for Evaluating Program Impacts**

Method Description

Quasi-experiments with constructed controls

Quasi-experimental designs involve comparing the performance of poverty

rates in regions targeted by the small enterprises to similar regions that have

not received assistance. Assignment to the two groups is non-random.

Rather, a comparison group is constructed after the fact. To the extent that

the two groups are similar, observed differences can be attributed to the

program with a high degree of confidence. Valid comparisons require that

the two groups be similar in terms of their composition with respect to key

characteristics, exposure to external events and trends, and propensity for

program participation. There are several types of designs that fall within this

general category, including simple difference-in-difference and regression

discontinuity.

Experiments with random assignment

The gold standard in impact assessment is experimental design with random

assignment to treatment and control groups. The critical element of this

design is randomization. Random assignment is an evaluation technique whereby people eligible for a program are assigned randomly to either of two groups — one that is subject to the program and allowed to receive program services (the treatment or program group) and one that is neither subject to the program nor allowed to receive program services (the control group). The purpose of random assignment is to ensure that any differences between the two groups, other than being subject to the program and receiving program services, is due to chance alone. (MDRC)

Random in this case does not mean haphazard; care needs to be taken to ensure that every eligible region has an equal chance of being selected for either group. Random assignment helps guarantee that the two groups are similar in aggregate, and that any extraneous factors that influence outcomes are present in both groups. For example, random assignment helps ensure that both groups of regions are similar in terms of the proportion of firms that are inherently more receptive to making needed changes in business practices, or that fluctuations in market conditions affect both groups equally. As such, the control group serves as the ideal counterfactual. Because of this comparability, claims that observed differences in outcomes between the two groups are the direct result of the program are more difficult to refute.

**Control groups**

There are ethical issues in using control groups, especially for the group denied services, assuming that the services are useful. “An examination of the methodologic issues of assessing evidence to inform policies and programs on the social determinants of health argues that randomized controlled trials (RCTs) are inadequate because randomization in studies of social factors is not always ethical, feasible, desirable or useful. An emphasis on evidence-based medicine has devalued information gained from sources other than RTCs, even in nonclinical research. In addition to its logistical challenges, randomly assigning participants in studies of social factors is problematic because it can create ethical issues and lead to a lack of generalizability. The health effects of social factors can be difficult to document in a controlled fashion, especially for “upstream” factors such as income and education, which influence health in complex ways over a long period of time. Interaction between upstream and downstream factors (such as physical activity), further complicates measuring their effects. RCT therefore cannot be treated as the "gold standard" for evidence.”

# ([Braveman PA](http://www.rwjf.org/vulnerablepopulations/search.jsp?author=Braveman%20PA), [Egerter SA](http://www.rwjf.org/vulnerablepopulations/search.jsp?author=Egerter%20SA), [Woolf SH](http://www.rwjf.org/vulnerablepopulations/search.jsp?author=Woolf%20SH) and [Marks JS](http://www.rwjf.org/vulnerablepopulations/search.jsp?author=Marks%20JS); When Do We Know Enough to Recommend Action on the Social Determinants of Health? [American Journal of Preventive Medicine (Supplement)](http://www.rwjf.org/vulnerablepopulations/search.jsp?pubtitle=American%20Journal%20of%20Preventive%20Medicine%20%28Supplement%29), [40(1)](http://www.rwjf.org/pr/search.jsp?pubtitle=American%20Journal%20of%20Preventive%20Medicine%20%28Supplement%29&volume=40&number=1), pp.S58-S66; **January 2011;**Robert Wood Johnson Foundation; <http://www.rwjf.org/vulnerablepopulations/product.jsp?id=71589>)

# Clinical trials

In clinical development, **clinical trials** are usually conducted to collect data for evaluation of the efficacy and safety of test treatments under investigation. To provide accurate and reliable assessment, well-controlled clinical trials are necessarily conducted. Clinical trial process is a lengthy and costly process, which is necessary to ensure a fair and reliable assessment of the test treatment under investigation. Clinical trial process consists of protocol development, trial conduct, data collection, statistical analysis and reporting. In practice, however, some controversial issues are commonly encountered regardless the compliance of good clinical practice.

These issues include, but are not limited to,

(1) appropriate of hypotheses for drug evaluation,

(2) the instability of classical sample size calculation,

(3) integrity of randomization/blinding,

(4) clinical strategy for endpoint selection,

(5) impact of protocol amendments,

(6) issue of multiplicity in clinical trials,

(7) feasibility of adaptive design methods,

(8) independence of data monitoring committee (DMC), (9) probability of success of clinical development, and

(10) other commonly encountered controversial issues, such as who gets to participate in the trials, whether various racial, ethnic, and gender groups are adequately represented, and ethical issues of those who do not benefit from the medicine.

(Derived from Controversial Issues in Clinical Trials; Mailman School of Public Health; Columbia University; <http://www.mailman.columbia.edu/events/display?site=msph&event_id=4436>)

A major way to **evaluate results is on the basis of evidence**. There is a hierarchy of desirable evidence:

* Category I:  Evidence from at least one properly randomized controlled trial. (But see discussion elsewhere in this book on the problems with randomized trials in humanitarian work.)
* Category II-1: Evidence from well-designed controlled trials without randomization.
* Category II-2: Evidence from well-designed cohort or case-control analytic studies, preferably from more than one center or research group.
* Category II-3: Evidence from multiple times series with or without intervention or dramatic results in uncontrolled experiments such as the results of the introduction of penicillin treatment in the 1940s.
* Category III: Opinions of respected authorities, based on clinical experience, descriptive studies and case reports, or reports of expert committees.

(Harris, R.P. et al. (2001). Current methods of the U.S. Preventive Services Task Force: a review of the process. American Journal of Preventive Medicine. April 20 (3 Supplement): 21-35)

Participant judgment and expert opinion.

This approach relies on people who are familiar with the intervention to make judgments concerning its impact. This can involve program participants or independent experts. In either case, individuals are asked to estimate the extent to which performance was enhanced as a result of the program – in effect, to compare their current performance to what would have happened in the absence of the program. While this approach is quite common, it is fraught with problems. It requires people to be able to determine the net effect of the intervention based solely on their own knowledge without reference to explicit comparisons. However, it may be the only option available given data and budget constraints.

(Nexus Associates, Inc; Assessing the Poverty Impact of Small Enterprise Initiatives;Prepared for: Working Group for Impact Measurement and Performance Committee of Donor Agencies for Small Enterprise Development; 8 September 2003)

### Originally introduced in the United Kingdom by the statutory public sector duties, impact assessments refer to the process by which every policy, procedure, practice, plan and strategy of an organization (including higher education institutions) is systematically reviewed and evaluated to ensure that they are not discriminatory and that they are making a positive contribution to equality. This is done by assessing how the impact they have differs (if at all) for different equality target groups. This normally requires the collection of statistical data for the area under review, and its analysis according to equality variables. It may also be necessary to consult and involve different groups of staff and service users, in order to obtain qualitative data.

The assessment process normally has two stages:

* initial screening for equality impact, possibly followed by
* more in-depth fullequality impact assessment, where there is a significant or important potential effect on equality that must be examined.

**Institutional memory**

Is there good reporting that will ensure that a project can be “remembered” and that this “memory” can be used as a learning tool?

Have terms of reference, findings, lessons learned and recommendations been placed in the public domain?

**What if the evaluation shows failure?**

**Admitting errors and failures**

Humanitarian organizations and other non-profits don’t like to admit errors and failures. Abraham Lincoln said, “Men are greedy to publish the successes of [their] efforts, but meanly shy as to publishing the failures of men. Men are ruined by this one sided practice of concealment of blunders and failures.”

However, the Robert Wood Johnson Foundation Evidence Matters project recently stated, “Unwelcome as the news might sometimes be, there is as much to learn from foundation programs that fail to meet expectations as there is from programs that have met or exceeded them—perhaps more. The truth is that not every program will succeed as planned. It is not surprising then that some Foundation-funded programs which seemed like good ideas at the time have not turned out to be so in practice.” (December 2010)

It is said in the high tech start up world that a person who has never experienced failure is not worth considering for investment of venture capital. “Failure is a necessary by-product of innovation…To understand failure, it is important to analyze its causes. Two obvious causes of failure are failure of premise or concept (that is, a bad idea), and failure of implementation (that is, a poorly executed idea)…two other factors…often go unmentioned. The first is power dynamics (*e.g.*, political influences, fiscal realities, leadership changes), which bear so heavily on an initiative’s success or failure that planners ignore or discount them at their own peril.

The second factor is an **institution’s capacity for self-analysis**. The effectiveness of an innovation can be undermined or even destroyed by an organization’s inability to be self-critical and open to reflection…People have a natural tendency to proclaim their successes and hide their failures. Funders like successful organizations with strong track records. This can result in organizations trumpeting (and recycling) their success stories.

The incentives to learn from their failures are less obvious.” (Greg Berman, Phillip Bowen, and Adam Mansky; Trial and Error: Failure and innovation in Criminal Justice Reform; Summer 2007; Executive Exchange; [www.courtinnovation.org/\_uploads/documents/failure.pdf](http://www.courtinnovation.org/_uploads/documents/failure.pdf))

**Abandoning the project.**

“Foundations don’t want to fund unsuccessful programs, but they may not

wish to abandon a project when first efforts fail.” (Let’s Discuss Evaluation; FSG Social Impact Advisors and James Irvine Foundation; Evaluation Toolkit for Trustees; June 2009; http://www.fsg.org/Portals/0/Uploads/Documents/PDF/Talking\_About\_Evaluation.pdf?cpgn=WP%20DL%20-%20Eval%20Toolkit%20Talking%20About%20Evaluation)

**Attach explicit consequences**, such as:

Allocating more funds to the program areas that show positive evaluation results and less to those that make no progress.

Examining cost per outcome of different grantees and shifting grants to the best performers.

Linking staff bonuses to evaluation results.

**Should failure be publicized?**

**Has the organization avoided mission creep?**

Mission creep is the pressure for a non-profit organization to add new goals and new program areas without subtracting any old ones. Mission creep results in an inability to carry out programs well.

**Classic non-profit organization evaluation criterion:**

What percent of money received is spent on fund-raising and administration, and what percent on direct delivery of services?

**Consistency**

“A…pitfall that…NGOs may fall into is inconsistency in their programs and results. This…is often due to a lack of clear goals, action plan, and an evaluative framework.  If NGOs fail to document their impact and fail to focus on empirical best practices, their efficacy will be limited. However, it is often difficult to assess the impact of NGO programs as a result of poor data, unclear objectives, rapidly changing circumstances, and poor-quality evaluations…performance cannot be quantified by anecdotal and circumstantial evidence…it is important to have an evaluation system geared toward improving the implementation of global health programs. If evaluations detect that harm is being done, resources are being wasted, or that there are inefficiencies in the system, action must be taken to address these problems. Conversely, if evidence-based evaluations indicate success, then more resources should be allocated to the relevant programs and measures to scale-up the success.”

# (The Importance of Avoiding Pitfalls in Global Health & Development; Unite for Sight; <http://www.uniteforsight.org/pitfalls-in-development/>; accessed 6/12/2010)

**Program Integration.** Is the program integrated with other programs by other service providers and organizations? Evidence shows that“a combined approach to tackling poverty and disease -- that brings together work on water and sanitation, health, education, and nutrition/food security -- achieves better results for the world's poorest… integrated aid programmes result in more effective and lasting solutions. Drawing from the evidence gathered, the report makes the following recommendations to international institutions, politicians, donors, and their NGO partners:

Commit high-level political leadership and financial resources to integrated approaches, where they can create greater impact.

Work with communities to design, implement, and evaluate integrated projects and programmes.

Incorporate experience and expertise across different sectors and stakeholders into development programmes.

Fund projects that demonstrate integrated approaches at the community level to learn what works, improve national plans, and inform scale-up strategies.

Provide flexible funding to partner countries in support of national priorities and integrated programmes.”

(PATH (2011, September 14). Combined approach to tackling poverty and disease achieves better results for the world's poorest, report says. ScienceDaily. Retrieved May 2, 2012, from http://www.sciencedaily.com­ /releases/2011/09/110914100525.htm)

**Scale-Up Evaluation.** Scale-up evaluation means the stage of research during which a project analyzes whether an intervention is effective in producing improved outcomes for individuals with disabilities when implemented in a real-world setting. During this stage of research, a project tests the outcomes of an evidence-based intervention in different settings. It examines the challenges to successful replication of the intervention, and the circumstances and activities that contribute to successful adoption of the intervention in real-world settings. This stage of research may also include well- designed studies of an intervention that has been widely adopted in practice, but that lacks a sufficient evidence-base to demonstrate its effectiveness.

**Indicators**

Indicators are more specific statements that describe how outcomes are being accomplished — detailed examples that can be observed and that are logically tied to an outcome.

What are **strong indicators**?There are six key criteria that can be used to check indicators and ensure that they are strong:

1. Is the indicator a direct measure of the outcome? The indicator should be indicative of the outcome rather than a predictor of the outcome. It should mean that the outcome has been achieved.

2. Is there a logical link between the indicator and the outcome? If the outcome reflects a change in behavior, the indicator should also reflect a change in behavior.

3. Is the indicator measurable and observable? An indicator should point to information that is tangible in the real world. Can it be seen, heard or read?

4. Are the indicators valid and reliable measures of the outcomes? Validity means, “Are you measuring what you intended to measure?” You want your indicators to be the closest, most accurate proxy for the outcomes as possible. Reliability means, “Can you collect the indicator data in a consistent way?” Consider both when selecting indicators.

5. Is it reasonable to expect the agency can collect data on the indicators? Sometimes, the indicators you choose might require a lot of time and resources to collect, or they might be difficult to gain access to.

6. Are the indicators useful? Will they help you understand what is going on in the program and where improvements may be needed? Think ahead and imagine the end of the evaluation process. Will the audience care about the progress made on these indicators? Will it be meaningful, given the overall intention and context of the program or initiative?

(Organizational Research Services; organizationalresearch.com)

“**Considerations/Limitations.**In selecting performance indicators, evaluators and service providers need to consider:

*The relevance of potential measures to the mission/objective of the local program or national initiative.* Do process indicators reflect program strategies/activities identified in mission statements? Do outcome indicators cover objectives identified in mission statements? Do indicators capture the priorities at the community level?  
  
*The comprehensiveness of the set of measures.* Does the set of performance measures cover inputs, outputs, and service quality as well as outcomes and include relevant items of customer feedback?  
  
*The program's control over the factor being measured.* Does the program have influence/control over the outputs or outcomes measured by the indicator? If the program has only limited influence over the outputs or outcomes being measured, the indicator may not fairly reflect program performance.  
  
*The validity of the measure.* Do the proposed indicators reflect the range of outcomes the program hopes to affect? Are the data free from obvious reporting bias?  
  
*The reliability and accuracy of the measure.* Can indicators be operationally defined in a straightforward manner so that supporting data can be collected consistently over time, across data gatherers, and across communities? Do existing data sources meet these criteria?  
  
*The feasibility of collecting the data.* How much effort and money is required to generate each measure? Should a particularly costly measure be retained because it is perceived as critically important?”

(Evaluation Strategies for Human Services Programs: A Guide for Policymakers and Providers; Adele Harrell *with* Martha Burt, Harry Hatry; Shelli Rossman; Jeffrey Roth, and William Sabol; The Urban Institute; Washington, D.C; http://www.ojp.usdoj.gov/BJA/evaluation/guide/documents/evaluation\_strategies.html)

“Some **examples of commonly-used [human rights] Key Performance Indicators (KPIs**):

Process/inputs

• Percentage of employees trained in Code of Conduct

(including human rights)

• Number of safety walks held per business unit

• Comparison of zero-tolerance vs. limited tolerance issue

(supply chain program)

Outcome/impacts

• Number and breakdown of code violations (e.g. respectful

treatment, discrimination, collective bargaining, employee

relations, employee privacy, right to organise, working hours)

• Lost time injury rate

• Number of contracts cancelled due to incompatibility with

human rights standards.”

(Institute for Human Rights and Business; The “State of Play” of

Human Rights Due Diligence: Anticipating the Next Five Years; June 2010; www.institutehrb.org)

**Implementation Hints**

Use checklists.

Build in audit points.

**Meta-Evaluation**

“For accountability purposes, and for reasons of clarity, evaluators should make sure that the link between their findings and the evidence used is clear. Output data should be presented, whenever available, not only for accountability purposes, but also to allow the reader to make his or her own judgement as to the overall findings of the evaluator. The logic of the evaluator should be clear and convincing. At the same time, it must not be

assumed that evaluators are infallible – methodological clarity enhances the accountability of the evaluator to those being evaluated…A ‘typical’ complex emergency is characterised by:

• a collapse of state functions

• intra-state rather than inter-state conflict

• difficulty in differentiating between combatants and civilians

• violence directed towards civilians and civil structures

• fluidity of the situation on the ground

• a lack or absence of normal accountability mechanisms

• the potential and actual development of war economies

• the potential for humanitarian assistance to prolong the conflict

• a multiplicity of actors” (Alistair Hallam; Evaluating Humanitarian Assistance Programmes in Complex Emergencies; Good Practice Review; Relief and Rehabilitation Network; September 1998; ODI)

Did the project actions or personnel change as a result of the evaluation? Were any of the recommendations carried out? If not, then the evaluation might have been useless. A good evaluation is not just an exercise designed to fulfill artificial requirements.

**What worked?** We have found that often a client perceives something we have done as having worked.  This is often good enough evidence for practitioners—the client reported satisfaction.  But there is the accompanying question of whether it was what we did that achieved the result, or was it other similarly timed factors, or a random effect, or an outside influence?

**Pitfalls and Cautions in Evaluations [Problems in evaluations; evaluation problems]**

A number of pitfalls and cautions have been found in the accomplishment and consequences of some evaluations. Some not easily avoidable:

The first relates simply to length of reporting;

the second to the representativeness of evidence, and the ability to generalize;

the third to the depth required of evidence on field activities;

the fourth to the optimal process for the synthesis of field based information to produce an analysis of overall performance.

the fifth has to do with a situation where one’s project has been evaluated, but others have not. Then, criticisms and shortfalls are recorded for one’s project, and not for the others. The others may claim great success, but have never actually been professionally evaluated.

a sixth has to do with coordinating evaluations, when more than one funder is present. Evaluations are usually funded by funders. In this case, the funders and other entities desiring an evaluation should get together and decide who will do it, how it will be done, when, and who will pay for it. An agreement can be drafted, or, in the UK terminology, the “Terms of Reference” (TOR).

a seventh has to do with “the tension between evaluation methodologies that satisfy the requirements of donors and those that are most useful for organizational learning and decision-making. It was clear by the end of the discussions that many considered these two imperatives irreconcilable. A participant in the closing plenary made the perhaps radical suggestion that the path forward may be to simply separate these two kinds of processes, instead of continuing to pretend a single methodology can meet the needs of external and internal audiences.” (Andrew Blum; Improving the State-of-the-Art: The Peacebuilding Evaluation Project Evidence Summit; United States Institute of Peace; January 11, 2012; http://www.usip.org/publications/improving-the-state-the-art-the-peacebuilding-evaluation-project-evidence-summit)

Some of these risks can be resolved through trade-offs.

The use of case studies carries the inherent risk of non-representative evidence.

**Do not confuse and confound assessment with evaluation**. For example, one-page surveys to elicit participants’ opinions about specific program activities such as guest speakers or panel presentations are not the same as a rigorous, systematic evaluation of a program.

“One of the great problems we had to deal with in Aceh was the proliferation of conflicting questionnaires and performance indicators. They were developed by well meaning people clueless of the distortions they introduced and in all cases missed the main recovery drivers and what we were doing in relation to them.” (Bill Nicol, LinkedIn, 6/7/2013)

**Scarce resources are often devoted to the intervention rather than to documenting participants, processes, and outcomes.** (Building Engineering and Science Talent (BEST). (2004). A Bridge for All: Higher Education Design Principles to Broaden Participation I Science, Technology, Engineering and Mathematics. San Diego:

BEST, 17-18)

**“Longevity does not ipso facto mean that the program is successful**. Without designing and incorporating an evaluation component into an intervention from the outset, it is unlikely that the impact and texture of the program will ever be measured.” (BEST 2004).

“Defining Program Participation.

Evaluating the Relationship between Participation and Outcomes.

Defining the Unit of Analysis.

Tracking the services received by participants.

Developing common agreements among agencies on program goals and required components.

Documenting service delivery by multiple agencies.

Measuring effects of the service delivery system.

Differentiating services integration from service comprehensiveness.

*(*Evaluation Strategies for Human Services Programs: A Guide for Policymakers and Providers; Adele Harrell *with* Martha Burt, Harry Hatry; Shelli Rossman; Jeffrey Roth, and William Sabol; The Urban Institute; Washington, D.C; <http://www.ojp.usdoj.gov/BJA/evaluation/guide/documents/evaluation_strategies.html>)

**Evaluators must avoid falling for “Potemkin Villages.”** These are impressive facades or shows designed to hide an undesirable fact or condition. The term comes from Grigori Potemkin, who allegedly built fake villages along a route that Catherine the Great of Russia was to travel in the Crimea in 1787.

**Controversies in Evaluations**

“[A] central debate in philanthropic circles -- **the tension between what the Center for Effective Philanthropy calls "program impact" and "institutional effectiveness."** The former is the data that the Gates Foundation lives by: doses of vaccines distributed, lives saved, dollars and manpower deployed to attack a problem. The latter refers to the sustainability of the change produced -- and is harder to judge. Many experts have found that good intentions still produce unintended consequences. When a Gates-backed NGO hires all the best nurses in a poor country, for example, it heals the sick but also siphons human resources from the local health infrastructure. Scaling up global development contributions risks provoking "Dutch disease," when a sudden flood of dollars causes damaging inflation in local economies. Laurie Garrett of the Council on Foreign Relations questions major outsider investments, in which "the particular concerns of wealthy individuals and governments drive practically the entire global public-health effort." The choices made by large foundations like Gates can overemphasize particular diseases at the expense of other, no less troubling public-health problems.”

(Gatekeepers: Is giving away money -- and lots of it -- really the best way to change the world? Dayo Olopade; The American Prospect; September 13, 2010; <http://prospect.org/cs/articles?article=gatekeepers>)

**Bottom Line**

Professor Ken Wiltshire, head of public administration at the University of Queensland Business School in Australia, posed two key questions as challenges for each manager:

• **"If your organization were to be abolished, would it be missed?"**

**• "If your organization was privatized, would anyone invest in it?"**

**References and Resources [Resources Chapter] [Resources and References] [resource chapter] [references chapter]**

Church, Cheyanne and Mark Rogers.

Designing for Results: Integrating Monitoring & Evaluation in Conflict Transformation Programmes.

Washington, DC: Search for Common Ground, 2005.

http://www.sfcg.org/programmes/ilr/ilt\_manualpage.html

Coalition for Evidence-Based Policy

Rigorous Program Evaluations on a Budget:

How Low-Cost Randomized Controlled Trials Are Possible in Many Areas of Social Policy

March 2012

http://coalition4evidence.org/wordpress/wp-content/uploads/Rigorous-Program-Evaluations-on-a-Budget-March-2012.pdf

Dead Aid: Why Aid Is Not Working and How There Is a Better Way for Africa, by Dambisa Moyo

Elliot, Michael, Tamra Pearson D’Estree, and Sanda Kaufman.

“Evaluation as a Tool for Reflection.”

Beyond Intractability - More Constructive Approaches to Destructive Conflict.

Sept. 2003

http://beyondintractability.org/essay/Evaluation\_Reflection/?nid=1268.

Evaluation Strategies for Human Services Programs: A Guide for Policymakers  
and Providers; Adele Harrell *with* Martha Burt, Harry Hatry; Shelli Rossman; Jeffrey Roth, and William Sabol; The Urban Institute; Washington, D.C; <http://www.ojp.usdoj.gov/BJA/evaluation/guide/documents/evaluation_strategies.html>

Evaluating Development Co-operation Summary of Key Norms and

Standards 2nd ed. Development Assistance Committee, 2006. OECD DAC

Network on Development Evaluation.

<http://www.oecd.org/dataoecd/12/56/41612905.pdf>

European Commission, EuropeAid Co-operation Office; Result Oriented Monitoring Handbook; July 2009; <http://ec.europa.eu/europeaid/how/ensure-aid-effectiveness/documents/rom_handbook2009_en.pdf>

Patton, Michael Q. “Conceptualizing the Intervention: Alternatives

for Evaluating Theories of Change.” Utilization-focused Evaluation.

Thousand Oaks: Sage Publications, 2008. 333-37. Web. http://www.

sagepub.com/upm-data/22861\_Chapter\_10.pdf

Professionalising the Humanitarian Sector

The study which was undertaken by Catherine Russ of RedR and Dr. Peter Walker of the Feinstein International Centre comprehensively investigates the need, motivations and potential benefits for the creation of an international framework for humanitarian professionalisation and sets out 13 core recommendations for the international humanitarian community to take forward.

<http://www.elrha.org/professionalisation>

Evaluating humanitarian action using the OECD-DAC criteria: An ALNAP guide for humanitarian agencies; Overseas Development Institute; 2006; <http://www.daraint.org/sites/default/themes/acrylic/downloads/ALNAP_EHA_Guide.pdf>

Active Learning Network for Accountability and Performance in Humanitarian Action

Jame Irvine Foundation

FSG Social Impact Advisors

Let’s Discuss Evaluation

June 2009

http://www.fsg.org/Portals/0/Uploads/Documents/PDF/Talking\_About\_Evaluation.pdf?cpgn=WP%20DL%20-%20Eval%20Toolkit%20Talking%20About%20Evaluation

W.K. Kellogg Foundation, *W.K. Kellogg Foundation Evaluation Handbook*, 1998, pp. 16,

http://www.wkkf.org/knowledge-center/resources/2010/W-K-Kellogg-Foundation-Evaluation-Handbook.aspx

Draft PEPH Evaluation Metrics Manual; Partnerships for Environmental Public Health (PEPH); National Institute of Environmental Health Sciences; National Institutes of Health; <http://www.niehs.nih.gov/research/supported/programs/peph/materials/index.cfm>

Nexus Associates, Inc; Assessing the Poverty Impact of Small Enterprise Initiatives;Prepared for: Working Group for Impact Measurement and Performance Committee of Donor Agencies for Small Enterprise Development; 8 September 2003

["Comparative Cost-Effectiveness Analysis to Inform Policy in Developing Countries"](http://www.povertyactionlab.org/publication/cost-effectiveness" \t "_blank); Poverty Action Lab; MIT; Cost-Effectiveness Methodology; <http://www.povertyactionlab.org/policy-lessons/cost-effectiveness>

Real-time evaluations of humanitarian action

An ALNAP Guide: Pilot Version

John Cosgrave, Ben Ramalingam, and Tony Beck; Overseas Development Institute; 2009; <http://www.alnap.org/pool/files/rteguide.pdf>

Ross, Marc Howard. (2000) “’Good-Enough’ Isn’t So Bad: Thinking about

success and failure in ethnic conflict management” Peace and conflict:

Journal of Peace Psychology 6(1). pp. 27–47

Voluntary Organizations for Professional Evaluation (VOPEs): Learning from Africa, Americas, Asia, Australasia, Europe and Middle East; Better Evaluation; <http://betterevaluation.org/resources/overview/Rugh_Voluntary_Organizations>; accessed 6/24/2013; Rugh, J., and Segone, M., editors (2013). Voluntary Organizations for Professional Evaluation (VOPEs): Learning from Africa, Americas, Asia, Australasia, Europe and Middle East*.* EvalPartners, UNICEF, IOCE. <http://www.mymande.org/sites/default/files/UNICEF_NY_ECS_Book2_web.pdf>

Schmidt, R.E., J.B. Bell, and JW. Scanlon *(1979),* "Evaluability Assessment: Making Public Programs Work Better," *Human Services Monograph Series, 14: 4-5.* Washington, DC; and Wholey, Joseph S. *(1994),* "Assessing the Feasibility and Likely Usefulness of Evaluation." In Joseph S. Wholey, Harry P. Hatry, and Katherine E. Newcomer (eds.), *Handbook of Practical Evaluation, 15-39.* San Francisco: Jossey-Bass

Search for Common Ground. “Evaluation Guidelines.” July 2006.

<http://www.sfcg.org/sfcg/sfcg_evaluations.html>

STEPS Steps to Transforming Evaluation Practice for Social Change,

http://www.stepstoolkit.org/

Ulrich Villis, Pia Hardy, Thomas Lewis; Creating Social Impact:

Strategic Use of Resources in the Social Sector; December 2009; Boston Consulting Group; [www.bcg.com](http://www.bcg.com)

# Troubleshooting Guide for Solving Problems: Common Problems, Reflection Questions, and Support Tools;

Work Group for Community Health and Development at the University of Kansas; <http://ctb.ku.edu/en/solveproblem/Troubleshooting_Guide_10.aspx>

World Bank; Evaluating the Impact of Development Projects on Poverty:

A Handbook for Practitioners; Judy L. Baker; 2000; [http://siteresources.worldbank.org/INTISPMA/Resources/handbook.pdf](http://www.linkedin.com/e/3oya7g-griw0aso-3f/plh/http%3A%2F%2Fsiteresources.worldbank.org%2FINTISPMA%2FResources%2Fhandbook.pdf/5MFy/" \t "_blank)

An excellent and comprehensive guidebook on social program evaluation.

**Appendices**

**Example 1: [Appendix One Appendix 1]**

**The International Response to Conflict and Genocide: Lessons**

**from the Rwanda Experience**

The report was written by a team of 52 independent experts

Humanitarian action cannot substitute for political action.

Coordination among UN agencies and between UN, bilateral and

NGO agencies was inadequate. Collaboration and coordination

between UN agencies was affected by overlapping mandates,

unclear division of responsibilities, and a regrettable rivalry.

The performance of NGOs was very mixed. A number delivered

high quality services and behaved professionally. But others

performed in an unprofessional and irresponsible manner that

resulted not only in duplication and wasted resources but, in

a few egregious cases, in unnecessary loss of lives. The NGO

community - and the donors who fund most - of it must

establish a Code of Conduct - and ensure adherence to it.

Preparedness and contingency planning was weak and there is a

need for a integrated humanitarian early warning system as

well as willingness of donors to provide up-front funding for

preparedness.

Military contingents from OECD countries have played

increasingly significant roles in support of humanitarian

operations. The report raises questions about the

predictability, effectiveness, costs and ability of the

military to participate collaboratively in humanitarian

operations. The role of the military needs to be thoroughly

assessed.

The international media played a mixed role in the Rwanda

crisis. While the media were a major factor in generating

worldwide humanitarian relief support for refugees, distorted

reporting on events leading to the genocide itself was a

contributing factor to the failure of the international

community to take more effective action to stem the genocide.

The distorted reporting reflected inadequate knowledge of

Rwandese culture and politics as well as low editorial

priority. The report recommends that the media conducts its

own self critical evaluation of the adequacy and impartiality

of its reporting of complex emergencies in the developing

world and draw lessons for more responsible reporting.

(Steering Committee For Joint Evaluation Of Emergency

Assistance To Rwanda; 12 March 1996; UN Department

Humanitarian Affairs Integrated Regional Information Network; <http://www.africaaction.org/docs96/rwan9604.2.htm>; accessed 1/18/2010)

**Example 2:**

Kristian Stokke; Humanitarian Response to Natural Disasters : **A Synthesis of Evaluation Findings**; 2007; Norwegian Agency for Development Cooperation; <http://www.norad.no/en/Tools+and+publications/Publications/Publication+page?key=109644>; accessed 1/18/2010)

**Description**

This report was commissioned by the Evaluation Department of Norad, which requested a synthesis report on the lessons learned in the field of humanitarian response to natural disasters. The report is intended to provide input to ongoing processes, including the preventive efforts of the Norwegian Ministry of Foreign Affairs, and to provide basic information for interested actors in the humanitarian sector.

The scale of Western humanitarian response has grown significantly since the end of the Cold War. This growth has been followed by a rising number of evaluations of humanitarian operations and meta-analyses of the international humanitarian system. However, the evaluations and lessons learned have a tendency to remain fragmented and case-specific, despite their many similarities as regards both the themes that are emphasised in the evaluations and the weaknesses that are exposed in the humanitarian response.

Evaluations of humanitarian response to natural disasters have many features in common. This applies primarily to the themes that are emphasised, but also to the lessons learned within these thematic areas. This report underscores five general themes in the available evaluation reports. In brief, they can be summarised as the need for and experience of: (1) linking of relief, recovery and development, (2) mapping and monitoring needs and target groups, (3) synergy between local, national and international capacities, (4) coordination of humanitarian actors and projects, and (5) disaster preparedness and vulnerability reduction.

The findings of this report can be summarised in five points: • Firstly, it is consistently found that the divide between humanitarian disaster response and development cooperation continues to prevail despite increased emphasis on the need to link relief, recovery and development. • Secondly, there is strong focus on needs assessment as a prerequisite for effective, equitable humanitarian response, but it is also commonly observed that actual assessment practice shows substantial deviations from this norm. • Thirdly, it is a common experience that international humanitarian response undermines rather than bolsters local capacity, despite growing emphasis on the importance of local capacity in humanitarian response and long-term vulnerability reduction. • Fourthly, it is a recurrent theme in evaluation reports that there is a great and persistent need to find effective mechanisms for coordinating the multitude of actors in the humanitarian system. • Fifthly, the evaluation reports show that there is growing awareness of the need for disaster preparedness and vulnerability reduction, but there are relatively few examples of good practices.

These evaluation findings show a clear need to reform the international humanitarian system to promote both synergies between international actors as coordinating facilitators and local vulnerability reduction and response capacity. The challenge for Norway as a donor lies not only in contributing to increased response capacity through improved funding and coordination, but also in helping to put in place strategies that focus on long-term vulnerability reduction and local response capacity. This indicates a need for selective choices of channels on the basis of three overarching objectives: (1) effective humanitarian relief, (2) rehabilitation, livelihood development and vulnerability reduction, and (3) coordination of relief, rehabilitation and livelihood development.

**Example 3:**

CARE France asked Groupe URD to conduct the **final evaluation of a project** which aimed to ‘improve the standard of living in 6 rural communities in the town of Corralillo based on sustainable development’. The 5-year project was co-financed by the European Commission and run jointly by CARE and the Asociación Cubana de Producción Animal (ACPA). It concerned an area affected by hurricanes.

The overall objective of the evaluation was:

- To evaluate the implementation and development of the project strategy, methodology and results and the key factors involved.

The specific objectives were to evaluate:

- the relevance/appropriateness of the project in relation to the context and the needs of the population;

- the effectiveness of the project and the degree to which it achieved its objectives;

- the efficiency of the project – that there is a balance between project inputs and outputs;

- the technical, economic, social and organisational impact of the project;

- the participation of beneficiaries and their degree of satisfaction;

- the viability of the results obtained;

- the appropriateness of technical and productive strategy and choices;

- the participation of the different local actors and the degree to which they fulfilled their obligations;

- the inclusion of gender considerations.

(Groupe Urgence Réhabilitation Développement Groupe URD ; <http://www.urd.org/en/news/general/news_evaluation.htm>)

**Example 4:**

# How Companies Can Help in Haiti

# Harvard Business Review Online

# http://blogs.hbr.org/cs/2010/01/how\_companies\_can\_help\_in\_hait.html

11:56 AM Thursday January 14, 2010   
by Timothy Ogden

The [news and images from Haiti](http://http://www.nytimes.com/2010/01/15/world/americas/15haiti.html?hp) are heartbreaking. Many individuals and companies are reaching for their wallets; others are wondering what they can do to help. But the best advice for corporate givers trying to figure out how best to respond is the old adage: "Don't just do something, stand there." To understand why, take a look back at the responses to other recent disasters. There is a [discernable pattern](http://www.alnap.org/pool/files/ALNAPLessonsEarthquakes.pdf), and not a good one:

1. Donations spike in the immediate aftermath.
2. A huge portion of the funds donated are spent on setting up disaster-relief operations that are no longer the primary need.
3. A flood of cash and materials cause a [logistics nightmare](http://http://www.foxnews.com/story/0,2933,583017,00.html) leading to waste and ineffectiveness, if not corruption.
4. Six months later, reconstruction stalls because the world's attention has moved elsewhere.
5. And, finally, a series of reports bemoan the fact that too many funds are devoted to disaster relief and not enough to disaster preparedness and reconstruction.

Companies are in the perfect position to break this cycle: They have the ability (and the obligation) to be thoughtful and strategic about how they handle their charitable giving. Here are a few ways businesses can help — and some principles everyone can apply to post-disaster giving:

**Don't earmark your donations for Haiti.** Funds for disaster relief are absolutely necessary in the short term — but immediate relief efforts are just one part of a long recovery process. By the time money earmarked for disaster relief arrives in charities' bank accounts for a particular disaster, recovery workers have already moved on to the much harder, much more expensive rebuilding phase. Rather than earmark a gift for Haitian disaster relief, direct your donations toward replenishing the cash and materials that disaster-relief agencies will expend in the next few weeks in Haiti, so they will be ready to respond immediately to the next disaster.

**Go with experience.** If you feel that you must give to disaster relief in Haiti, make sure you are giving to an organization that has extensive experience in Haiti and people already on the ground. They will be much more effective because of their existing knowledge of communities, cultural norms, and power dynamics.

**Give money.** Gifts-in-kind may seem like an appealing and useful way to contribute but they tend to cause huge logistical problems that dramatically undermine their value. Money gives those responding to the disaster the ability to act flexibly, according to the needs at hand. If you do have materials you are convinced will be useful (construction supplies, computers, and so on) ask the charity your working with whether it can effectively use what you can give. (And be prepared to hear that they'd rather have the cash.)

**Look ahead.** Long after their immediate health and safety needs are taken care of and the media spotlight has moved on, Haitians will still benefit from your organization's help in rebuilding their lives and livelihoods. One way to do this — and engage employees and customers — is to match the dollars they contribute for immediate relief with a corporate gift for reconstruction, to be given in six or eight months. By that time it will be clear which areas of the rebuilding effort are underfunded. You'll also have time to thoroughly vet agencies, projects, and so forth, to ensure that your donations will do the most good.

Timothy Ogden is the president of [*Sona Partners*](http://www.sonapartners.com) and the editor-in-chief of [*Philanthropy Action*](http://www.philanthropyaction.com) , a journal for high-net-worth donors. He was formerly chief knowledge officer at [*Geneva Global*](http://www.genevaglobal.com), an international philanthropy advisory company.

**Example 5**

**Example of Good Coordination**

The Lancet, [Volume 360, Issue 9329](http://www.thelancet.com/journals/lancet/issue/vol360no9329/PIIS0140-6736%2800%29X0308-X), Pages 335 - 336, 27 July 2002

<[Previous Article](http://www.thelancet.com/journals/lancet/article/PIIS0140-6736%2802%2909559-4/fulltext)|[Next Article](http://www.thelancet.com/journals/lancet/article/PIIS0140-6736%2802%2909523-5/fulltext)>

doi:10.1016/S0140-6736(02)09574-0[clear](http://www.thelancet.com/popup?fileName=cite-using-doi)[Cite or Link Using DOI](http://www.thelancet.com/popup?fileName=cite-using-doi" \t "newWin)

# Coordination of humanitarian aid—a UN perspective

Original Text

[Erick de Mul](http://www.thelancet.com/search/results?fieldName=Authors&searchTerm=Erick+de%20Mul) [a](http://www.thelancet.com/journals/lancet/article/PIIS0140-6736%2802%2909574-0/fulltext?version=printerFriendly" \l "aff1)

Médecins Sans Frontières feeding centre in Kuito, Bié Province, 2002

The *Lancet's* June 22 Editorial[1](http://www.thelancet.com/journals/lancet/article/PIIS0140-6736%2802%2909574-0/fulltext?version=printerFriendly" \l "bib1) calling on the UN to put its humanitarian house in order and encouraging non-governmental organisations (NGOs) to support this process drew attention to one of the most important problems facing the international community today. Where states have failed to fulfil their responsibilities, or in countries where warring parties are incapable of fulfilling, or worse, indifferent to their responsibilities to protect and care for civilian populations, how best can humanitarian agencies save people whose lives are in the balance?

Hundreds of individual agencies around the world are now working in humanitarian assistance. The efforts of most of these organisations are remarkable. Millions of lives are saved every year by programmes that are launched, often in extreme circumstances, by UN agencies and NGOs. The resources available for humanitarian action, however, are almost never enough.

For every former Yugoslavia, where sufficient aid funding was available, there are many more for-gotten emergencies, such as Angola, where agencies are working around the clock to save millions of people with only a fraction of the resources they need. Ensuring that resources are used in the most effective way possible, that hundreds of agencies are focused on the core problems, and that the people who are really in trouble get help requires strong, action-oriented coordination.

Coordination of humanitarian action is a time-consuming and sometimes difficult process. Organisations come into coordination structures voluntarily. They have different, occasionally contradictory mandates. Variations in approach, perspective, and priorities are inevitable and reflect the diversity of the humanitarian community.

When disagreements occur, the lead coordination agency must make every effort to find common ground between partners. The purpose of coordination structures is to keep humanitarian operations moving forward in order to avoid paralysing the flow of aid. There are too many lives at stake in complex emergencies to allow differences and disagreements to divert resources and attention away from the real problems. Differences exist within any humanitarian operation, but an operation that is effective and coordinated is an operation in which differences are dealt with, sometimes publicly, sometimes not, but in a way that aims at getting as much humanitarian assistance as possible to the people who need it the most.

The humanitarian operation in Angola is widely regarded as one of most effectively coordinated in the world. Ten UN agencies, 100 international NGOs, and more than 300 national organisations are working in consultation with 11 government departments and ministries and in all 18 Angolan provinces. The Coordination structure in Angola has been built up over a 10-year period. All facets of humanitarian programmes are organised by this network. At the strategic level, humanitarian partners discuss and agree overall priorities. At the operational level, partners collaborate to ensure the smooth and effective functioning of programmes on the ground. At the sectoral level, partners set goals and objectives and establish common approaches for specific sectors, including food security, health, nutrition, and water and sanitation.

Hundreds of humanitarian organisations are working right now in partnership in Angola to expand operations to meet the emergency needs of millions of people. When the Memorandum of Understanding, which ended hostilities, was signed by the Angolan government and UNITA (National Union for the Total Independence of Angola) on April 4, humanitarian operations were already underway for two million people. Assistance was being delivered to more than 170 separate locations, and agencies were stretched to the limit, trying to cope with the massive internal displacement that occurred during the final stages of the war.

During the past 3 months, in a major, all-out effort to respond to the emergency, agencies have systematically pushed into areas where they were unable to go during the war. The results are impressive.

Multisectoral operations covering food assistance, health, nutrition, water, sanitation, and protection have been launched in 22 new locations and programmes are underway in 31 of the 35 areas where UNITA families are being quartered. UN agencies and NGOs are spreading out across the country, trying to reach populations who are at serious risk. However, landmine infestation, the appalling state of Angola's infrastructure, and lack of funding are major constraints and have set the pace of the operation. In spite of these difficulties, partners are unified by the primary need to continue until everybody in jeopardy has been reached.

In the specific case of emergency food assistance, the World Food Programme (WFP) has begun feeding programmes in more than 50 new locations since April. As soon as the Angolan government authorised UN assistance for families in the UNITA quartering sites, WFP began releasing food to the few NGOs, including Médecins Sans Frontières, who were willing to establish emergency supplementary feeding programmes in the quartering areas ahead of the security clearances required by the UN and most NGOs. Such cooperative action saved lives and prevented further deterioration of the already precarious situation in many of these sites.

To reach people in need since the end of hostilities, WFP has overcome huge logistical challenges associated with transporting food to extremely isolated areas that lack even the most basic infrastructure. People have been reached with food despite impassable roads, broken bridges, and mined airstrips. This work is continuing. Specialised logistics teams are operating at break-neck speed throughout the country to open transport routes for humanitarian assistance.

With every humanitarian agency in the country operating at full capacity, cooperation and burden-sharing is the only way forward. WFP has had to take over the direct distribution of food in several locations where NGO partners do not have sufficient capacity. WFP is also coordinating with the government of Angola and other agencies that have their own food pipelines. Such cooperation allows every agency to focus their resources where they can be used to best effect.

The rapid and systematic expansion of the humanitarian operation in Angola has been possible only because of the highly effective coordination mechanisms that exist between the UN, NGOs, and the Angolan government. The challenges facing the coordination structure in Angola are immense. The UN now estimates that up to three million Angolans are in desperate need of humanitarian assistance.

Meeting these needs, however, will be impossible unless sufficient funding is provided for humanitarian operations. In the case of Angola, the lack of financial support is startling. Only 30% of the UN consolidated interagency appeal for 2002 has been funded to date, which means that there is an enormous gap between the needs that must be met and the resources that are available to humanitarian organisations.

Reform of the UN is important, and serious efforts aimed at strengthening the system and making it more effective are needed. But we must avoid the cynicism of careless criticism and recognise that real concrete progress has been made in recent years. Humanitarian assistance has been improved and pragmatic coordination mechanisms have ensured that hundreds of aid workers are aiming for the same objective, saving the lives of people who would otherwise suffer or die. We must not mistake serious resource and political constraints for bureaucratic inertia. Humanitarian organisations in Angola are working together towards the same objective—ensuring that the peace that has finally come to the country after decades of warfare lasts, and that the Angolan people who have suffered so much survive and prosper with dignity.

## References

[1](http://www.thelancet.com/journals/lancet/article/PIIS0140-6736%2802%2909574-0/fulltext?version=printerFriendly#back-bib1) Editorial. Beyond trading insults in international humanitarian aid. Lancet 2002; 359: 2125.

[a](http://www.thelancet.com/journals/lancet/article/PIIS0140-6736%2802%2909574-0/fulltext?version=printerFriendly#back-aff1) UN OCHA, UN Office for the Coordination of Humanitarian Affairs, Angola

E de Mul is the UN Humanitarian Coordinator in Angola

**Appendix Six**

**Why Evaluations Fail**

**“Smart Transportation Emission Reduction Strategies"**

( [http://www.vtpi.org/ster.pdf](http://www.vtpi.org/ster.pdf" \t "_blank) )  
This report investigates the optimal (best overall, taking into account all benefits and costs) transportation emission reduction strategies. Current evaluation methods tend to undervalue mobility management (also called Transportation Demand Management or TDM) strategies that increase transport system efficiency by changing travel behavior, due to biases that include (1) ignorance about these strategies; (2) failure to consider co-benefits; (3) failure to consider rebound effects of increased fuel economy; (4) belief that mobility management impacts are difficult to predict; (5) belief that mobility management programs are difficult to implement; and (6) belief that vehicle travel reductions harm consumers and the economy. More comprehensive and objective analysis tends to rank mobility management strategies among the most cost-effective emission reduction options. This report describes ways to correct current planning bias so mobility management solutions can be implemented to the degree optimal.”

Todd Litman; Victoria Transport Policy Institute

**Appendix Seven:**

**Effective human service systems**

Effective human service systems deliver services that are:

**Integrated and Coordinated.** Effective service systems are set up in a way that allows individuals and families to easily access services through single points of entry and makes sure they do not “fall through the cracks.” The Urban Institute provides the following description of the difference between service integration and coordination/collaboration: "Integration is characterized by features such as common intake and 'seamless' service delivery, where the client may receive a range of services from different programs without repeated registration procedures, waiting periods, or

other administrative barriers. In contrast, coordinated systems generally involve multiple agencies providing services, but clients may have to visit different locations and re-register for each program to obtain services.” Best practices include:

· Common intake forms and applications

· Co-location of services, where public agencies and nonprofit providers share space, or, at minimum, assign staff from several agencies to serve shared clients in one place.

In addition to making services more “user-friendly” for community members, service integration and coordination approaches allow staff of different agencies to become more knowledgeable about resources available. Further, combining resources such as staff and facilities helps reduce duplication and may allow participating agencies to stretch their funding further.

**Holistic and Client-Centered.** In a holistic system, the components of that system do not just address an isolated need or problem. Staff in different program or service areas understand how needs are inter-related, actively work to identify other issues and needs, and address them through referrals and partnerships… The components of the system work together so that no one service area says, “that’s not my problem.”

**Delivered by trained, qualified, and committed staff**. Effective systems have human resources policies and practices that ensure that all staff have the qualifications, experience, interest, and training to perform their jobs effectively. Government agencies and nonprofits alike should be connected to and ensure their staff receive training from the appropriate professional associations for their field. Staff also need to be committed to serving diverse communities, this includes spending the extra time often needed to find creative solutions to the complex challenges low-income and foreign-born populations in particular face, rather than giving up and saying “this is somebody else’s problem.”

**Accessible.** Effective service systems assess the barrier to access across service areas and target populations. They then ensure that various different services are designed and delivered in a way that ensures they are accessible to those who most need them. This is especially critical for low-income families, limited English proficient people, and people with disabilities. Best practices include:

· Locate services near neighborhoods where target communities live or within walking distance of public transportation when selecting sites for new programs, classes, or services.Ensure an accessible right of way, with good signage in various local languages.

· Provide transportation assistance, for example through stipends or tokens, ride shares, volunteers, etc…

· Ensure services are accessible to parents of young children by providing child care.

Numerous studies of family literacy and English language instruction programs, for example, find that on-site child care is critical to success. Providers can offer stipends to participants to pay their own babysitters, pay stipends to community members to provide child care on site during classes, or use volunteers to provide child care during classes, for example. Note that high school students can often receive service learning credit for this type of volunteer work.

For a variety of adult education and training programs, some national organizations and state agencies have identified the provision of support services--including child care and transportation--as a standard for quality programs. For example, the State of Maryland’s Adult English Language Program includes as a standard: “The program provides access to a variety of services, directly or through referrals to cooperating agencies: child care,

transportation, health services, employment counseling, educational counseling, legal advice, personal and family counseling, assessment of learning disabilities and other handicapping conditions, and translators and interpreters.”

**Well managed**. Effective service systems need skilled leadership. In *well-run* programs and services, classes and appointments happen on time, staff and volunteers follow through, questions get answered, and the goals of the program--for improving lives and communities--are met. This requires good planning, clearly defined outcomes, and skilled managers who ensure that programs and services have the resources they need to meet their objectives, including appropriate data systems and technology.

**Identify and respect clear rights for the people being served**. A growing number of human services systems and individual service providers are developing a client or consumer *bill of rights*. These are often brief fact sheets that define client or consumer rights--such as a right to confidentiality, a right to respectful treatment, a right to not be

discriminated against, and a right to a grievance procedure if the client or participant feels s/he has been mistreated or not received quality services. Some of these are required by law in the United States for recipients of federal financial assistance, such as the right to nondiscrimination.

**Designed and planned with the involvement of the people being served**. Effective service systems provide clear and consistent opportunities for the people being served to provide input into how programs and services are designed and delivered. They will have insights about *what it will take to make a program work* that the service provider won’t know--such as the times of day that would be most convenient. A growing number of human service fields encourage the active participation of people who they hope will benefit from programs and services in planning them. This has been part of a *consumer movement* in which the people who receive services--often called *consumers--*have demanded the right to have a say in the design, decision making, and delivery of programs and services. Best practices include:

· Creating consumer advisory groups. Fields such as health care, mental health, and disability services have led the way in creating such structures for consumer involvement. Some federal funding sources require consumer involvement in the form of public participation and involvement formal community planning councils and/or advisory boards, committees, or commissions. Such requirements exist in transportation planning, for example.

· Conducting surveys, focus groups and holding community meetings to consult with and involve the community, including people and organizations that represent diverse racial and ethnic groups, religious groups, income groups, age groups, gender, disability, and other perspectives.

**Consistently evaluated and shown to have positive outcomes.** Effective service systems have policies and practices to support consistent system-wide evaluation of performance and outcomes. The system should provide the resources, training, and tools needed to carry out evaluation. Approaches to evaluation should also be flexible and ensure that program participants are included in defining “success.” Again, it goes back to asking community members what is important to them and what it would take for them

to be satisfied with both delivery of service and the final results. Evaluation results should be used to improve programs and ensure they are high quality.

Derived from Hilary Binder Aviles, Mosaica: The Nonprofit Center for Development and Pluralism with contributions from Maryland Association of Nonprofit Organizations; A Report on the Best Practices Research of the Capacity Building Initiative; Howard County, Maryland; September 2009; Howard County Department of Citizen Services

**Appendix Nine**

# How to Live with – and Benefit from – Evaluation Requirements

By Elizabeth Duffrin

Institute for Comprehensive Community Development

LISC  
September 10, 2010

http://www.instituteccd.org/library/1546

“…Funders can expect results too quickly…They may ask for a more extensive evaluation than you can afford. And they tend to change their minds about what they want you to do. “They can start asking questions that have to do with where their heads are at rather than what was originally agreed to.”

…Conducting evaluation on your own terms strengthens both your program and the relationship you're building with program officers at a foundation…

**Negotiate.** Sit down with your program officers at the start and find out what they think is credible evidence of success. But make sure to educate them about the challenges you face so that their expectations match what you can accomplish.

For example, the goal of your project might be to reduce unemployment in the neighborhood. There are factors outside your control impacting that statistic, so you might suggest evaluating outcomes that you can influence more directly, such as the number of residents attending your job training program and meeting regularly with an employment counselor. “Figure out what you can do and then sell that as reasonable.”

**Serve yourself.** The ultimate purpose of evaluation should be to improve your own work. So even if a funder isn't interested in all of your goals, make sure the evaluation answers your questions, too. Above all, avoid leaving the direction of your project entirely up to your funders. Foundations frequently change priorities, “and you have no control over that,”…“If you focus on your funders, you will lose your way.”

**Start early.** Evaluation can seem so daunting that it often gets left until the end of a project. “Evaluating your progress from the start allows you to make mid-course corrections—and that’s the real point of evaluation,”…

If your goal is far off, look for short-term indicators that tell you if you're headed in the right direction. If you want to reduce crime in the neighborhood, how many residents are turning out  for community meetings on the topic? How many are joining the Parent Patrol or the Neighborhood Watch? If the numbers are low, why and what can you do about it?  
When data collection becomes a routine part of your job, evaluation is less of a burden. For example, neighborhood residents might not mind filling out a quick survey at the end of each community meeting…

**Don't leave it to the experts.** Even if you hire outside evaluators, you will still need to supervise them to make sure your questions get answered. “Lots of times people hand [evaluation] off because they find the process annoying or time consuming,”…But don't,…“The consequences of it going badly or getting off track are real.”

Make sure the evaluators you hire are skilled both in research and in communicating that research to the outside world—whether to community residents, policy makers or foundations. “You want someone able  to listen and explain to people rather than retreating into their expert world and language,”…“If they do that, you’ve got to fire them.”

**Involve the community.** Community residents can be a huge resource for your evaluation. If your researcher only has time to interview 20 people, training residents to conduct interviews could significantly expand that sample size. “Lots of people find it interesting to be part of [the evaluation]—checking in with the neighbors or visiting the police station or interviewing the local businesses.”

Just make sure to keep the tasks manageable, such as conducting simple interviews with five neighbors. And make sure topics are appropriate for resident evaluators. Interviews about recreation, traffic or safety are fine, while questioning ex-offenders about arrest and probation is probably not. Leave questions that are too technical or sensitive to the experts.

**Go beyond anecdotes.** Telling human stories about your success is important, but don't forget to back them up with numbers, which tend to carry more weight in the foundation board room. You may already have more data than you realize, said Dewar, as anecdotes collected over time can become data. “Sometimes you're sitting on data and because you're a narrative-style person, you don't realize what evidence is [already] inside the story.”

If your project involves engaging different segments of the community in a development project, who is attending your meetings? How many residents, how many business owners? Who is speaking up? Does the nature of the comments change over time? For example, do you hear people specifically recognizing another interest while stating their own opinion? How often does that happen? Looking for things to count can provide a backbone to your narrative. Be creative.  
Be honest. If your project falls short of its goals, be open about that with your funder. Scrutinize your work and describe what you learned and what you might attempt next.  Describe the challenges you encountered.

The societal problems that non-profits tackle are extremely complex, and this is a chance to educate your funder in a way that can deepen your long-term working relationship, Dewar explained. Besides, foundations are so accustomed to “huge shortfalls between rhetoric about what you were going to do and what is produced,” that submitting a candid and thorough evaluation is unlikely to hurt you…

**Appendix Ten**

**Going Wrong**

(London) October 19, 2010 -- The Ethiopian government is using development aid to suppress political dissent by conditioning access to essential government programs on support for the ruling party, Human Rights Watch said in a report released today. Human Rights Watch urged foreign donors to ensure that their aid is used in an accountable and transparent manner and does not support political repression.   
  
The 105-page report, "[Development without Freedom: How Aid Underwrites Repression in Ethiopia](http://lists.hrea.org/phplist/lt.php?id=Kh5TAwdeCAFQBxgGCwQGWk4DBFMMBA%3D%3D" \t "_blank)," documents the ways in which the Ethiopian government uses donor-supported resources and aid as a tool to consolidate the power of the ruling Ethiopian People's Revolutionary Democratic Front (EPRDF).   
  
"The Ethiopian government is routinely using access to aid as a weapon to control people and crush dissent," said Rona Peligal, Africa director at Human Rights Watch. "If you don't play the ruling party's game, you get shut out. Yet foreign donors are rewarding this behavior with ever-larger sums of development aid."   
  
Ethiopia is one of the world's largest recipients of development aid, more than US$3 billion in 2008 alone. The World Bank and donor nations provide direct support to district governments in Ethiopia for basic services such as health, education, agriculture, and water, and support a "food-for-work" program for some of the country's poorest people. The European Union, the United States, the United Kingdom, and Germany are the largest bilateral donors.   
  
Local officials routinely deny government support to opposition supporters and civil society activists, including rural residents in desperate need of food aid. Foreign aid-funded "capacity-building" programs to improve skills that would aid the country's development are used by the government to indoctrinate school children in party ideology, intimidate teachers, and purge the civil service of people with independent political views…  
  
Farmers described being denied access to agricultural assistance, micro-loans, seeds, and fertilizers because they did not support the ruling party. As one farmer in Amhara region told Human Rights Watch, " leaders have publicly declared that they will single out opposition members, and those identified as such will be denied ‘privileges.' By that they mean that access to fertilizers, ‘safety net' and even emergency aid will be denied."   
  
Rural villagers reported that many families of opposition members were barred from participation in the food-for-work or "safety net" program, which supports 7 million of Ethiopia's most vulnerable citizens. Scores of opposition members who were denied services by local officials throughout the country reported the same response from ruling party and government officials when they complained: "Ask your own party for help."   
  
Human Rights Watch also documented how high school students, teachers, and civil servants were forced to attend indoctrination sessions on ruling party ideology as part of the capacity-building program funded by foreign governments. Attendees at training sessions reported that they were intimidated and threatened if they did not join the ruling party. Superiors told teachers that ruling party membership was a condition for promotion and training opportunities. Education, especially schools and teacher training, is also heavily supported by donor funds.   
  
"By dominating government at all levels, the ruling party controls all the aid programs," Peligal said. "Without effective, independent monitoring, international aid will continue to be abused to consolidate a repressive single-party state."   
  
In 2005, the World Bank and other donors suspended direct budget support to the Ethiopian government following a post-election crackdown on demonstrators that left 200 people dead, 30,000 detained, and dozens of opposition leaders in jail. At the time, donors expressed fears of "political capture" of donor funds by the ruling party.   
  
Yet aid was soon resumed under a new program, "Protection of Basic Services," that channeled money directly to district governments. These district governments, like the federal administration, are under ruling party control, yet are harder to monitor and more directly involved in day-to-day repression of the population.   
  
During this period the Ethiopian government has steadily closed political space, harassed independent journalists and civil society activists into silence or exile, and violated the rights to freedom of association and expression. A new law on civil society activity, passed in 2009, bars nongovernmental organizations from working on issues related to human rights, good governance, and conflict resolution if they receive more than 10 percent of their funding from foreign sources.   
  
"The few independent organizations that monitored human rights have been eviscerated by government harassment and a pernicious new civil society law," Peligal said. "But these groups are badly needed to ensure aid is not misused."   
  
As Ethiopia's human rights situation has worsened, donors have ramped up assistance. Between 2004 and 2008, international development aid to Ethiopia doubled. According to Ethiopian government data, the country is making strong progress on reducing poverty, and donors are pleased to support Ethiopia's progress toward the United Nations Millennium Development Goals. Yet the price of that progress has been high.   
  
When Human Rights Watch presented its findings to donor officials, many privately acknowledged the worsening human rights situation and the ruling party's growing authoritarian rule. Donor officials from a dozen Western government agencies told Human Rights Watch that they were aware of allegations that donor-supported programs were being used for political repression, but they had no way of knowing the extent of such abuse. In Ethiopia, most monitoring of donor programs is a joint effort alongside Ethiopian government officials.   
  
Yet few donors have been willing to raise their concerns publicly over the possible misuse of their taxpayers' funds. In a desk study and an official response to Human Rights Watch, the donor consortium Development Assistance Group stated that their monitoring mechanisms showed that their programs were working well and that aid was not being "distorted." But no donors have carried out credible, independent investigations into the problem…

"In their eagerness to show progress in Ethiopia, aid officials are shutting their eyes to the repression lurking behind the official statistics," Peligal said. "Donors who finance the Ethiopian state need to wake up to the fact that some of their aid is contributing to human rights abuses."

**Appendix Eleven**

**“THE FIVE CONDITIONS OF COLLECTIVE SUCCESS**

Our research shows that successful collective impact initiatives typically have five conditions that together produce true alignment and lead to powerful results: a common agenda, shared measurement systems, mutually reinforcing activities, continuous communication, and backbone support organizations.

***Common Agenda*** Collective impact requires all participants to have a shared vision for change, one that includes a common understanding of the problem and a joint approach to solving it through agreed upon actions. Take a close look at any group of funders and nonprofits that believe they are working on the same social issue, and you quickly find that it is often not the same issue at all. Each organization often has a slightly different definition of the problem and the ultimate goal. These differences are easily ignored when organizations work independently on isolated initiatives, yet these differences splinter the efforts and undermine the impact of the field as a whole. Collective impact requires that these differences be discussed and resolved. Every participant need not agree with every other participant on all dimensions of the problem. In fact, disagreements continue to divide participants in all of our examples of collective impact. All participants must agree, however, on the primary goals for the collective impact initiative as a whole…

Funders can play an important role in getting organizations to act in concert…

***Shared Measurement Systems*** Developing a shared measurement system is essential to collective impact. Agreement on a common agenda is illusory without agreement on the ways success will be measured and reported. Collecting data and measuring results consistently on a short list of indicators at the community level and across all participating organizations not only ensures that all efforts remain aligned, it also enables the participants to hold each other accountable and learn from each other’s successes and failures.

It may seem impossible to evaluate hundreds of different organizations on the same set of measures. Yet recent advances in Web-based technologies have enabled common systems for reporting performance and measuring outcomes. These systems increase efficiency and reduce cost. They can also improve the quality and credibility of the data collected, increase effectiveness by enabling grantees to learn from each other’s performance, and document the progress of the field as a whole…

***Mutually Reinforcing Activities*** Collective impact initiatives depend on a diverse group of stakeholders working together, not by requiring that all participants do the same thing, but by encouraging each participant to undertake the specific set of activities at which it excels in a way that supports and is coordinated with the actions of others.

The power of collective action comes not from the sheer number of participants or the uniformity of their efforts, but from the coordination of their differentiated activities through a mutually reinforcing plan of action. Each stakeholder’s efforts must fit into an overarching plan if their combined efforts are to succeed. The multiple causes of social problems, and the components of their solutions, are interdependent. They cannot be addressed by uncoordinated actions among isolated organizations…

***Continuous Communication*** Developing trust among nonprofits, corporations, and government agencies is a monumental challenge. Participants need several years of regular meetings to build up enough experience with each other to recognize and appreciate the common motivation behind their different efforts. They need time to see that their own interests will be treated fairly, and that decisions will be made on the basis of objective evidence and the best possible solution to the problem, not to favor the priorities of one organization over another.

Even the process of creating a common vocabulary takes time, and it is an essential prerequisite to developing shared measurement systems. All the collective impact initiatives we have studied held monthly or even biweekly in-person meetings among the organizations’ CEO-level leaders. Skipping meetings or sending lower-level delegates was not acceptable. Most of the meetings were supported by external facilitators and followed a structured agenda…

***Backbone Support Organizations*** Creating and managing collective impact requires a separate organization and staff with a very specific set of skills to serve as the backbone for the entire initiative. Coordination takes time, and none of the participating organizations has any to spare. The expectation that collaboration can occur without a supporting infrastructure is one of the most frequent reasons why it fails.

The backbone organization requires a dedicated staff separate from the participating organizations who can plan, manage, and support the initiative through ongoing facilitation, technology and communications support, data collection and reporting, and handling the myriad logistical and administrative details needed for the initiative to function smoothly. Strive has simplified the initial staffing requirements for a backbone organization to three roles: project manager, data manager, and facilitator.

Collective impact also requires a highly structured process that leads to effective decision making… adapt for the social sector the Six Sigma process that GE uses for its own continuous quality improvement. The Strive Six Sigma process includes training, tools, and resources that each SSN uses to define its common agenda, shared measures, and plan of action, supported by Strive facilitators to guide the process.

In the best of circumstances, these backbone organizations embody the principles of adaptive leadership: the ability to focus people’s attention and create a sense of urgency, the skill to apply pressure to stakeholders without overwhelming them, the competence to frame issues in a way that presents opportunities as well as difficulties, and the strength to mediate conflict among stakeholders.”

(John Kania and Mark Kramer; Collective Impact: Large-scale social change requires broad cross-sector coordination, yet the social sector remains focused on the isolated intervention of individual organizations; Winter 2011; Stanford Social Innovation Review; <http://www.ssireview.org/articles/entry/collective_impact/>)

**Appendix Twelve**

**Australian Aid Evaluation: Independent Review of Aid Effectiveness**

# Terms of reference

[Background](http://www.aidreview.gov.au/termsofreference/index.html#background)  
[Objective](http://www.aidreview.gov.au/termsofreference/index.html#objective)  
[Scope](http://www.aidreview.gov.au/termsofreference/index.html#scope)  
[Management arrangements](http://www.aidreview.gov.au/termsofreference/index.html#management)  
[Approach](http://www.aidreview.gov.au/termsofreference/index.html#approach)  
[Timing](http://www.aidreview.gov.au/termsofreference/index.html#timing)

## Background

The Australian aid program aims to assist developing countries reduce poverty and achieve sustainable development, in line with Australia's national interests. The program has doubled in size over the last five years to an estimated $4.3 billion in 2010–11 and, on current economic projections, will double again to meet the Government's commitment to increase Australia's aid to 0.5 per cent of gross national income by 2015–16.

The Government, Parliament and taxpayers need to be confident that this significant investment is both effective and efficient in fulfilling its objectives.

The Government has taken a range of measures to improve the effectiveness and efficiency of the aid program. The Office of Development Effectiveness (ODE), which was established in 2006, has completed a number of reviews and evaluations of the program, including the Annual Review of Development Effectiveness, which is tabled in Parliament. The findings of this analysis are used to improve aid program planning and implementation. AusAID, the lead agency within the Government on the aid program, has rigorous systems and processes in place to ensure that the aid program is well managed and prioritised. These systems are reviewed and improved regularly. A review of advisers engaged under the aid program is currently being conducted jointly with developing country partners, and a review of procurement and agreements processes has commenced. An audit of the aid program by the Australian National Audit Office in 2009 found that AusAID had effectively managed the increases in the program up to that time.

To ensure that the further increase in the aid budget to 2015–16 is well managed and meets the Government's objectives, a review of the aid program will be conducted. This will be the first independent public review of the aid program commissioned by the Australian Government since the Simons Review in 1996.

This review will draw on the experience of the last five years and relevant international experience and make recommendations regarding the structure of the program and the planning, implementation and review arrangements needed to support delivery of a substantially enlarged aid investment. This review will make a strong aid program even better.

## Objective

To examine the effectiveness and efficiency of the Australian aid program and make recommendations to improve its structure and delivery.

## Scope

In particular the review will focus on:

1. The structure of the program, noting in particular:
   * the appropriate geographic focus of the program, taking into account partner country absorptive capacities
   * the appropriate sectoral focus of the program, taking into account Australia's area of comparative advantage and measured development effectiveness results
   * the relative focus of the aid program on low and middle-income countries
   * the relative costs and benefits of the different forms of aid, including the role of non-government organisations and the appropriate balance between multilateral and bilateral aid funding arrangements.
2. The performance of the aid program and lessons learned from Australia's approach to aid effectiveness.
3. An examination of the program's approach to efficiency and effectiveness and whether the current systems, policies and procedures in place maximise effectiveness.
4. The appropriate future organisational structure for the aid program, including:
   * AusAID's organisational structure for aid delivery
   * arrangements for the coordination of ODA across the public service
   * coordination of Australia's ODA with other donors and institutions.
5. The appropriateness of current arrangements for:
   * review and evaluation of the aid program, including an examination of the role of the Office of Development Effectiveness and options to strengthen the evaluation of the aid program
   * the management of fraud and risk in the aid program.

The review will involve an examination of broader international thinking on aid effectiveness and will draw on work by the OECD DAC (including the most recent peer review of the Australian aid program), work on the approach and experience of non-state donors (such as the Clinton and Gates Foundations and non-government organisations) and the range of audits undertaken by the ANAO.

## Management arrangements

The review will be conducted by a panel consisting of:

* Mr Sandy Hollway, AO,(Chairman), former secretary of two Australian Government departments and CEO of the Sydney Organising Committee for the Olympic Games, previously, an official of the Australian Department of Foreign Affairs and Trade for 16 years, an Australian diplomat at four overseas posts and Head of the International Division and Deputy Secretary of the Department of Prime Minister and Cabinet with responsibility for foreign aid and other international matters;
* Dr Stephen Howes, Director, International and Development Economics, Crawford School of Economics and Government at the ANU, previously worked as the Lead Economist for India for the World Bank and as Chief Economist at AusAID;
* Ms Margaret Reid, AO, has extensive experience with Australian non-government organisations (NGOs) working in international aid as the former President of the Executive Committee of the Australian Council for International Development. Ms Reid is also the first female President of the Australian Senate, and former World President of the Commonwealth Parliamentary Association;
* Mr Bill Farmer is a former senior diplomat. Mr Farmer was Head of Mission in the two largest recipients of Australian aid, Indonesia and Papua New Guinea. Mr Farmer was also the Deputy Permanent Representative of Australia to the United Nations;
* Mr John Denton, CEO and Partner at Corrs Chambers Westgarth is a Prime Ministerial appointee to the APEC Business Advisory Council and a member of the Boards of the Business Council of Australia and the Commonwealth Business Council. He has firsthand experience of development and conflict through postings to Bangladesh and Iraq. He is Chairman of Australia for United Nations High Commissioner for Refugees.

The panel will be supported by a secretariat led by AusAID and drawn from a range of other Government agencies, including central agencies and agencies involved in the delivery of the aid program, as necessary. The panel will draw on expert advice as required.

## Approach

The review will consult extensively across the Australian Government, non-government organisations and other key stakeholders in the Australian community. Fieldwork will be conducted to consult with a selection of Australia's bilateral and multilateral partners.

## Timing

The review will commence in November 2010 and be completed by April 2011.

Last reviewed: 20 December, 2010

**Appendix Thirteen**

# Inquiry to assess the effectiveness of the government's PREVENT strategy, UK

# <http://iengage.org.uk/component/content/article/1-news/455-inquiry-to-assess-the-effectiveness-of-the-governments-prevent-strategy>

# 7/24/2009

# Accessed 2/3/2011

# The Communities and Local Government Committee of parliament is to undertake an [inquiry](http://www.parliament.uk/parliamentary_committees/clg/clgpn090721pve.cfm" \t "_blank) into ‘Prevent’, the Government’s strategy on Preventing Violent Extremism.

The inquiry is currently inviting written submissions on the effectiveness to date of Prevent, and its likely effectiveness in the future, with a focus on these questions:

1. Is the Prevent programme the right way of addressing the problem of violent extremism, or are there better ways of doing it?
2. How robust is the Government's analysis of the factors which lead people to become involved in violent extremism? Is the 'Prevent' programme appropriately targeted to address the most important of those factors?
3. How appropriate, and how effective, is the Government's strategy for engaging with communities? Has the Government been speaking to the right people? Has its programme reached those at whom it is—or should be—aimed?
4. Is the necessary advice and expertise available to local authorities on how to implement and evaluate the programme?
5. Are the objectives of the 'Prevent' agenda being communicated effectively to those at whom it is aimed?
6. Is the Government seeking, and obtaining, appropriate advice on how to achieve the goals of the 'Prevent' programme?
7. How effectively has the Government evaluated the effectiveness of the programme and the value for money which is being obtained from it? Have reactions to the programme been adequately gauged?
8. Is there adequate differentiation between what should be achieved through the Prevent programme and the priorities that concern related, but distinct, policy frameworks such as cohesion and integration?

Such an inquiry is long overdue and Muslims are strongly encouraged to contribute submissions to assist the committee in their appraisal of the strategy’s objectives, methods and results. This is your opportunity to move beyond mere criticism of the Government’s policies and instead feed into the constructive process of evaluating how the Prevent strategy has been implemented and whether the methods of implementation have had their desired effect.

Many of you will have knowledge of how the millions of pounds of PVE money has been spent locally and whether it was distributed fairly or not. ENGAGE urges everyone, individuals and Muslim organisations to contribute submissions - whether detailed or short - to the inquiry. ENGAGE has written [previously](http://iengage.org.uk/component/content/article/1-news/222-engage-exclusive-qf-memo-to-advisors-on-todays-damaging-times-story" \t "_blank) about how government funds have been spent on funding discredited organisations such as the Quilliam Foundation.  
  
The deadline for submissions is September 17, 2009. Further details and guidelines in making submission can be found [here](http://www.parliament.uk/parliamentary_committees/clg/clgpn090721pve.cfm" \t "_blank).

**Appendix Fourteen**

**Sample for Evaluating Merit of Projects**

**National Endowment for the Arts, US, Creative Placemaking**

The following are considered during the review of Statements of Interest and, in more detail, during the review of applications:

The **merit** of the project, which includes the

* Potential of the project to achieve results consistent with the NEA outcome for Livability: Strengthening communities through the arts. This includes the potential:
  + To contribute to the livability of the community and integrate design and the arts into the fabric of community life.
  + For lasting impact and, as appropriate, the ability to serve as a model for other communities.
  + Impact on the community, artists, design professionals, and arts organizations.
* Appropriateness of the proposed performance measurements and their ability to provide evidence that the NEA Livability outcome was achieved.
* Quality of the proposed partnership, including the required partners, and engagement of the private and public sectors in support of the project.
* Appropriateness of the project to the distinct character and qualities of the community.
* Extent to which the project engages the public in planning for and participating in the project.
* Appropriateness of the project to the organization's mission.
* As appropriate, plans for documentation, evaluation, and dissemination of the project results.
* Ability to carry out the project based on such factors as the:
  + Appropriateness of the budget and its community funding support and sustainability strategy.
  + Quality and clarity of the project goals and design.
  + Resources involved.
  + Qualifications of the project's personnel.
  + Readiness to meet requirements of the National Environmental Policy Act and/or the National Historic Preservation Act, where relevant.
  + Likelihood that the project will be completed within the proposed period of support.
* Where appropriate, potential to reach underserved populations such as those whose opportunities to experience the arts are limited by geography, ethnicity, economics, or disability.

**Appendix Fifteen**

**Evaluating Planning**

“General Plan quality is commonly recognized as an integral part of effective plan implementation (Baer, 1997; Burby & May et al., 1997). Evaluations of plan quality range from identifying characteristics of a high quality plan (Baer, 1997; Kent 1964) to case studies that evaluate plan policies as they relate to supporting sustainability (Berke & Conroy, 2000). Baer (pp. 338-339) outlines the following critical components to consider when preparing a General Plan; employment of these components enhances the quality of the plan and can ease implementation.

**•** Adequacy of context: Explain the context and purpose of the plan.

**•** “Rational Model” considerations: Show basic planning considerations

based on underlying theory and measurable criteria.

**•** Procedural validity: Explain the process for creating the plan.

**•** Adequacy of scope: Demonstrate that the plan is comprehensive in nature

and connected to issues and concerns outside of planning––for example,

education.

**•** Guidance for implementation: Document implementation tools and

procedures.

**•** Approach, data, and methodology: Document data sources and identify

relevance to policies.

**•** Quality of communication: Ensure document is well prepared and ideas

clearly presented.

**•** Plan format: Ensure the document is correctly formatted, well organized,

and easy to use.

**Characteristics of Effective General Plan Policies**

1. Effective policy should be explicit and directive; if not mandatory.

2. Effective policy should entail incentives that make it likely to be

implemented.

3. Effective policy should be clearly expressed, understandable and

accessible to those who must implement it or are affected by it.

4. Effective policy should be based on *and* make explicit reference to a

substantial factual basis (e.g. a technical study, data base or model.

5. Effective policy should be explicitly linked to performance standards or

indicators enabling the policy’s results to be monitored…

Baer [William Baer published an article in 1997 in the *APA Journal* that has…become a classic reference on General Plan evaluation] sets out general criteria for plan assessment. The titles (below) are followed by lists of criteria that implement their intent:

1. Adequacy of context – plans are not self-evident; explain the context to the public.

2. Rational model considerations – show the underlying theory and its criteria in the plan.

3. Procedural validity – what went on in making the plan; who was involved?

4. Adequacy of scope – the plan orientation to other jurisdictions and the

world.

5. Guidance for implementation – most plans do something. What

implementation tools?

6. Approach, data, and methodology – where did the data come from; how

was it used?”

(Richard W. Lee, Paul Wack, Judy Deertrack, Scott Duiven, Lisa Wise; The California General Plan Process and Sustainable Transportation Planning;

Mineta Transportation Institute, College of Business, San José State University; MTI Report 01-18; MAY 2002)

**Appendix Sixteen**

**Global Reporting Initiative (GRI)**

# GRI Reporting: 11 Principles = Sustainability reporting guidelines of the Global Reporting Initiative (GRI), published in 2002. This section outlines the GRI's 11 principles, covering the economic, environmental and ethical indicators of a company's performance. Seven main areas of activity: governance (Engaging), our market (Accessing), health (Protecting), the environment (Sustaining), our workplace (Balancing), and community involvement (Involving). The report also includes reviews of each of our operating businesses in the UK, Ireland, Germany and the Isle of Man. Five-Part Materiality Test developed by AccountAbility that we applied as a desk-based research tool to our last year's report. Starting from financial year 2005/06 we will further implement this newly developed method with the aim to better inform our decisions. Community (BITC) sustainability reporting guidelines, the Association of British Insurers' (ABI) Social Disclosure Guidelines and LBG community investment reporting.

The GRI sets out 11 Principles towards which reporting organisations should strive. We report against these to ensure that we present a balanced and reasonable account of our economic, environmental and social performance. Throughout the report we cross-reference the information in the report to the relevant GRI indicators. We also provide a GRI index on this website.

1. Transparency

To generate trust with those who are affected by our operations we need to be transparent about how we do business. Reporting forms an important part of this activity.

Prior to this year's Report we commissioned research that resulted in a method for us to use that would recognise both the views of stakeholders and our ability to manage corporate responsibility issues. This newly developed method helps us map corporate responsibility issues for O2 that fits within our overall risk management process and that can be applied through an online system.

The content of this year Corporate Responsibility report was based on the findings of our initial application of this new method in addition to other more conventional sources of information, such as opinion leader research and media reviews.

Our Report is externally verified and is one of the media we use to inform individuals and organisations that are affected by our operations. Increasingly, we are including more information on our website to enable easy access. The additional references to BITC, ABI and LGB reporting guidelines aim to ensure that stakeholders can make informed judgements about our operations.

2. Inclusiveness

We believe it is important to capture the opinions of a wide range of stakeholders. The decision what to report on was made by O2's reporting team based on a number of information sources. Our new method - that combines stakeholder feedback with how we manage the various social, environmental and ethical issues that we face as a business - and other stakeholder input, such as media reports and opinion leader research, helped us prioritise issues to be included in this Corporate Responsibility report. We have undertaken a number of research initiatives to gauge the opinions of employees, investors, non-governmental organisations, business associates, the general public and the media. Our assurance providers tested the scope and balance of the information contained in the Report against a selection of external media sources and outputs of internal and external stakeholder engagement.

In future we aim to record stakeholder feedback systematically through an online management tool. This should help us ensure that as many O2 people as possible understand the needs of our stakeholders.

This Report is aimed at a diverse audience and is compiled in response to feedback from diverse stakeholders. The detailed table of contents/sitemap of the report guides the readers from a number of entry points to find what it interesting and relevant to them.

3. Auditability

The GRI states that "data and information should be recorded, compiled, analysed and disclosed in a way that would enable internal auditors or external assurance providers to attest to its reliability". For the purpose of this report we engage with professional auditors with a strong track record of financial auditing and social, environmental and ethical assurance work. Our internal auditors also include corporate responsibility indicators in their audits. In addition, we engage external auditors and certification bodies in our operations to cover environmental management, health and safety, mast siting and quality of our operations. We are continuously improving our record keeping, supporting audit reviews and supply evidence on request.

Our environmental data is recorded at the six monthly O2 plc Environmental Forum meetings and electronically through an online system that has been reviewed by our internal audit department. We aim to improve the accuracy of our environmental reporting by including measurement as part of our service level agreements with our suppliers. Our human resources data is improving through new systems that have been designed with diversity metrics in mind.

For the purpose of this Report we asked our assurance providers to test statements made in the report and review the underlying processes and procedures for management of social, environmental and ethical matters. The auditors' observations are included in their assurance statement.  New for this year we had our community (Involving) section of the report and the underlying activities reviewed by a second assurance provider - the Corporate Citizenship Company.

4. Completeness

The GRI states that all information that is material to readers for assessing the reporting organisation's economic, environmental, and social performance should appear in the report in a manner consistent with the declared boundaries, scope, and time period. We aim to report on all social, environmental and ethical aspects of our business that we have a direct control over, excluding joint ventures or minority shareholdings. Increasingly, we also seek to report on areas that are beyond our direct control, but where we have influence, such as our distribution network. To portray a more holistic view of our operations we are engaging with our suppliers to widen the scope for our reporting, especially to the wider environmental impact of our supply chain. We report annual performance data on a wide range of issues and cover areas that may have a longer-term impact, such as child protection.

5. Relevance

As stated by the GRI - relevance relates to the threshold at which information becomes significant enough to be reported. This process is informed through our stakeholder engagement and research activities. This year we have also conducted a materiality assessment of the key social, environmental and ethical issues we face. We aim to build on this initial research to include systematic stakeholder feedback in our risk management tool.

6. Sustainability Context

We aim to embed sustainability/corporate responsibility into the fabric of our business, this is also true for our reporting practices. We aim to achieve this by reporting our operations against a wide context of issues, such as child protection and the economic benefits of our business to society. By adopting the LBG reporting tool we also try to capture the wider societal impacts of our community investment activity. We aim to demonstrate how the big issues influence and relate to our day-to-day business. Our plans to increase the systematic capture of stakeholder engagement should also add to our knowledge of the effect we have on society and the physical environment.

7. Accuracy

Accuracy of the Report is necessary to enable readers to make judgments based on its findings. Our internal external assurance processes as well as stakeholder feedback help provide for accurate reporting. We include some estimated figures, if exact data is not available, in our reporting, where it can aid the reader in forming an opinion of our performance. Any discrepancies in our data are fully explained in the Report. We are also committed to expanding our data measurement to cover more environmental data for our suppliers.

8. Neutrality

GRI states that "…reports should avoid bias in selection and presentation of information and strive to provide a balanced account of the reporting organisation's performance". We aim to present a balanced picture by reporting our performance clearly and openly. To accentuate the neutrality of this report we asked a number of independent individuals and organisations to offer their feedback, which we have included in full throughout this report. Our external auditors have also made an assessment of the information to ensure it is a balanced presentation of the key issues.

9. Comparability

GRI states that "the reporting organization should maintain consistency in the boundary and scope of its reports, disclose any changes, and re-state previously reported information". Since our formation in November 2001 we have worked to establish our data measurement procedures and practices. These have improved over the last three years and we believe our measurements now provide a more accurate picture of our operations. This should enable improved comparability of our performance in the future. We also engage in benchmarking activities, such as the LBG and participate in sustainability indices.

10. Clarity

To provide clarity, reports should be written and laid out in a way that enables them to be understood by a range of readers. We aim to improve the clarity of our reporting. The Plain English Commission has also assessed our Group report and we have included a section in German to reach our stakeholders in the German marketplace. We offer more information on our corporate responsibility practices on our website at [www.o2.com/cr](http://www.o2.com/cr" \o "Corporate responsibility practices (opens in a new window)" \t "_blank). The website has been built under the scrutiny of external accessibility assessors to enable easy access for those who use special accessibility equipment.

11. Timelines

We report our corporate responsibility performance annually through a dedicated report and through the O2 plc Annual Report and Financial Statements, the O2 plc Annual Review and our dedicated corporate responsibility website. We update our information on our website throughout the year.

**Appendix Seventeen**

**Example of a waste of money, time, and effort**

The Arc is an advocacy group headquartered in the United States, for people with intellectual and developmental disabilities.

# “The Arc Launches a New Brand Identity

Posted on [March 1, 2011](http://blog.thearc.org/2011/03/01/the-arc-launches-a-new-brand-identity/)

The Arc Blog

The Arc is proud to unveil a new brand identity! Starting here, starting now, you’ll find a dynamic new look for our blog, Website and publications. And along with that comes a comprehensive new brand identity which will serve to unite our chapters across the country under the banner “Achieve With Us.” It’s at once an invitation and a command, as it represents our movement’s determination to achieve even more in terms of participation, inclusion, acceptance and respect for people with I/DD in the years to come. In the logo’s design, you’ll see the energy and determination The Arc brings to supporting and embracing people with I/DD and their families across their lifetimes and across many diagnoses. You’ll notice a bold color combination of orange and yellow chosen specifically because it is unique in the realm of nonprofit organizations, and can be exclusively associated with The Arc much like a specific shade of pink is associated with breast cancer awareness.

Together, the elements of the brand represent a strong, energetic organization working to provide opportunity for people with I/DD; opportunity for hope, opportunity for growth and opportunity for change. We know a thing or two about change as we have been on the front lines of a movement that has ushered in dramatic changes in the law, in schools, in health care, in communities and often on a very personal level in individual’s lives. Check out this short video about The Arc and stay tuned for some compelling stories of individuals and families touched by The Arc as we introduce you to The Hage’s tomorrow, Adrian Forsythe on Thursday and Sam Jenkins on Friday. We think you’ll want to hear what they have to say.”

[**http://blog.thearc.org/2011/03/01/the-arc-launches-a-new-brand-identity/**](http://blog.thearc.org/2011/03/01/the-arc-launches-a-new-brand-identity/)

Questions to ask include, how much did this cost, how much time and energy did it involve, what will it accomplish in terms of actual delivery of services? Is this rebranding just a fad, keeping up with other organizations, or possibly done to compete with other disability advocacy groups?

**Appendix Eighteen**

Design for Success Program Checklist

Center for American Progress

Doing What Works

A. Are the goals clear and the cost estimates accurate?

1. Clear understanding of the problem in measurable terms

2. Clear goals expressed as outcomes over a deﬁned time scale

3. Accurate cost estimates

4. Suffcient funds to achieve goals

B. Is this the right approach?

5. Fills a clear gap in current government services

6. Cannot achieve goals through expansion or modiﬁcation of existing programs

7. Little duplication or overlap with current federal programs

8. Reasonable conclusion that this approach is the best option

9. Reasonable total cost for expected outcomes compared to alternatives

C. Why will the program work?

10. Reasonable expectation that key stakeholders will alter their ways of working or behavior

11. Professional, independent research indicates approach will work

12. Program expands on what worked on the smaller scale, or plan to test before rolling out on the large scale

D. Does the program establish the right incentives?

13. Establishes clear incentives that are aligned with program goals

14. Beneﬁciaries reasonably expected to change behavior

15. Beneﬁciaries have been consulted and indicate that the program is likely to work

16. Strong understanding of areas most prone to gaming, and risks mitigated in program design

E. How will the program be implemented

17. Reasonable timeline for program rollout

18. Implementers have bought into the program and are prepared to carry it out

19. Suffcient plan for hiring, training, and IT development

20. Reasonable procurement strategy and suffcient staff to negotiate and manage contracts

21. Design takes into account views from key outside stakeholders

22. Clear strategy for minimizing impact on negatively affected parties

23. Reasonable effort to minimize unintended consequences

24. Strong risk-mitigation strategy in place

F. How will you monitor success and rethink the approach

25. Key indicators of success identiﬁed for each goal

26. Clear plan for collecting timely, accurate data that reﬂects outcomes

27. Takes advantage of existing data systems where possible

28. Coordinates data collection with other programs where possible

29. Reasonable plan for rethinking approach

30. Low sunk costs

31. Can be quickly phased out or terminated if deemed ineffective or no longer necessary

**Appendix Nineteen**

**Three Question Evaluation**

“Kevin Starr, founder and director of the Mulago Foundation, said the nonprofit advocacy community sometimes deludes itself about the impact it’s having on social problem-solving. (Overly wordy, self-important and confusing mission statements, he said, are just one symptom of this behavior.)

Starr, who coaches nonprofits on how to be more effective, said he uses three questions to determine whether a nonprofit is doing the job. Do its leaders know clearly what it is trying to accomplish? “If you could only measure one thing to know if you fulfilled your mission, what would it be?” Starr asked. Do nonprofits measure the right things to determine success and do they measure them well enough? “Measuring well means taking an honest stab at understanding what change you’re creating,” he said. Is it needed? Does it work like it is supposed to? Will it get to those who need it and enough of those who do? Will people use it immediately?”

(Marcia Stepanek; Exploring Failure; November 3, 2010; Stanford Social Innovation Review; <http://www.ssireview.org/opinion/entry/exploring_failure>)

Mulago Foundation:

## “How we think about impact

We measure impact because it’s the only way to know whether our money is doing any good. In fact, we don’t invest in organizations that don’t measure impact – they’re flying blind and we would be too. Those organizations that do measure impact perform better and evolve faster, and discussions around measuring impact almost always lead to new ideas about effectiveness and efficiency.

Everyone’s got their own definition of impact and here’s ours: Impact is a change in the state of the world brought about by an intervention. It’s the final result of behaviors (outcomes) that are generated by activities (outputs) that are driven by resources (inputs).

We’re a small shop, so we needed to develop an approach with enough rigor to be believable, but simple enough to be doable. When we work with organizations we use these five steps to determine impact and calculate bang for the donor buck.

1. Figure out what you’re trying to accomplish: the real mission

You can’t think about impact until you know what you’re setting out accomplish. Most mission statements don’t help that much. We re-formulate the mission in a phrase of ~8 words or less that includes 1) a target population (or setting), 2) a verb, and 3) an ultimate outcome that implies something to measure – like this:

* Get African one-acre farmers out of poverty
* Prevent HIV infection in Brazil

If we can’t get to this kind of concise statement, we don’t go any further- either because they don’t really know what they’re trying to do or because we simply wouldn’t be able to know if they’re doing it.

2. Pick the right indicator

Try this: figure out the single best indicator that would demonstrate mission accomplished. Ignore the howls of protest: it’s a really useful exercise. Here are examples relating to the missions shown above:

* Change in farm income
* Decrease in HIV infection rates

Sometimes that best indicator is doable, and that’s great. Other times you might need to capture it with a carefully chosen – and minimal – combination of indicators. When there is a behavior with a well-documented connection to impact – like children sleeping under mosquito nets – you can measure that and use it as a proxy for impact. Projects that can’t at least identify a behavior to measure are too gauzy for us to consider. Notice that while things like “awareness” or “empowerment” might be critical to the process that drives behaviors, we’re interested in measuring the change that results from that behavior.

We don’t pretend that this method captures all of the useful impacts and accomplishments of a given organization and their intervention. What it does do for us as philanthropic investors is answer the most critical question: did they fulfill the mission?

3. Get real numbers

You need to 1) show a change and 2) have confidence that it’s real. This means that

* You got a baseline and measured again at the right interval, and
* You sampled enough of the right people (or trees, or whatever) in the right way.

There are two parts to figuring this out: the logical side and the technical side. With an adequate knowledge of the setting, you can do a lot by just eyeballing the evaluation plan - looking carefully at the methods to be used to see if they make sense. Most bad schemes have an obvious flaw on close examination: they didn’t get good baseline data, they ask the dads when they ought to ask the moms, they’re asking in a culturally inappropriate way. The technical part has mostly to do with sample size, and a competent statistician can easily help you figure what is adequate.

Measure \*Lasting impact. In most cases it’s also vital to measure impact at reasonable intervals to ensure that impacts are sustained. There aren’t any hard and fast rules about intervals; usually the nature of a given intervention suggests a suitable interval for follow-up.

4. Make the case for attribution

If you have real numbers that show impact, you need to make the case that it was your efforts that caused the change. This is the hardest part of measuring impact, because it asks you to be able to say what would have happened without you. When real numbers show there has been a change, a useful thing to ask is “what else could possibly explain the impact we observed?”

There are three levels – in ascending order of cost and complexity - of demonstrating attribution:

1. Narrative attribution: You’ve got before-and-after data showing a change and airtight story that shows that it is very unlikely that the change was from something else. This approach is vastly overused, but it can be valid when the change is big, tightly coupled with the intervention, involves few variables (factors that might have influenced the change), and you’ve got a deep knowledge of the setting.
2. Matched controls: At the outset of your work, you identified settings or populations similar enough to ones you work with to serve as valid comparisons. This works when there aren’t too many other variables, you can find good matches, and you can watch the process closely enough to know that significant unforeseen factors didn’t arise during the intervention period. This is rarely perfect; it’s often good enough.
3. Randomized controlled trials: RCT’s are the gold standard in most cases and are needed when the stakes are high and there are too many too many variables to be able to confidently say that your comparison groups are similar enough to show attribution.

5. Calculate bang-for-the-buck

Now that know you’ve got real impact, you need to know what it cost. You can always generate impact by spending a ton of money, but it won’t give good value for the philanthropic dollar and it won’t be scalable (and it probably won’t last). Stick with the key impact you’ve chosen; don’t get sucked in to the current trend of trying to monetize every social impact you can think of.

The easiest – and arguably most valid – way to calculate bang-for-the-buck is to divide the total donor money spent by the total impact. In organizations that do more than one kind of project, it is often possible to split out what they spent for their various impacts. Remember that start-ups are expensive and don’t worry so much about their current figures, but do see if their projections for steady-state operations make sense and assume (as we learned the hard way) that they are usually at the very optimistic end of the scale.

In the end though, the key to figuring out real impact is an honest, curious, and constructive skepticism. Everyone benefits from a rigorous look at impact: the doers, the funders, the social sector itself, and most importantly, those who are hoping for a brighter day ahead.”

(<http://www.mulagofoundation.org/ideas/r/how-we-think-about-impact>)

**Appendix Twenty**

**Self-Evaluation of an Organization that claims to be very effective**

BRAC

“[R]eaches 138 million of the poorest people in nine countries in Asia and Africa, and recently launched in Haiti. Its mission is to empower people and communities in situations of poverty, illiteracy, disease and social injustice. Its annual operating budget is half a billion dollars.”

“BRAC has created 8.5 million self-employment opportunities and made $5 billion in micro-loans to over 6 million borrowers. BRAC's schools graduated 3.8 million students from primary schools and 2.3 million from pre-primary schools, with nearly 1.8 million children currently enrolled in its 66,000 schools.”

“The idea was not to prove they had all the answers before they started, but to find out what worked and apply the lessons." From the beginning, BRAC let incompetent staff go and set up a training center for those who remained. As BRAC facilitated the development of small enterprises by the people they serve, BRAC sought and found ways to lower the costs of production in order to reduce the expenses of borrowing.”

“[I]nterviews with [Sal Giambanco](http://omidyar.net/team/sal-giambanco" \t "_new), Partner, Omidyar Network (a primary BRAC funder) and [Susan Davis](http://en.wikipedia.org/wiki/Susan_Davis_%28BRAC_USA%29" \t "_new), President & CEO, BRAC USA, reveal some of the key factors that make BRAC so uniquely effective in its holistic approach to poverty alleviation. BRAC has achieved so much because it:

1. Constantly tested, measured, and modified its solutions in order to maximize its effectiveness in solving poverty--from the very beginning. Never seeking to make its reports look better for funders, never assuming that it had the answers, always trying to find better solutions.
2. Engaged in the community to develop solutions together, not as an outside force thinking it had the answers and would drop resources onto the community and then leave.
3. Recognized the limits that government and religious and social norms placed on women, especially with regard to working, even though women were counted on to help support their families. Within those constraints, BRAC identified industries where women would be permitted to work (e.g., dairy), helped them set up microenterprises (through training and microlending), and also helped to find a way for the dairy products to reach the market.
4. Gradually developed a holistic approach in Bangladesh for three decades before replicating the model--first in Afghanistan 10 years ago. As BRAC developed through an iterative process, BRAC developed many interventions, coupling traditional microfinance loans with other programs like health care, education, human rights and legal services, and social enterprise. Its microfinance program targets the poorest of the poor, and adapts for the varied local contexts in which BRAC works.
5. Specifically designed programs--based on the concept of "massification"-- that could be expanded exponentially to serve the large populations of poverty-stricken people in Bangladesh. This positioned BRAC to be particularly effective in other regions with dense populations of people living in extreme poverty. As Abed said, "Small is beautiful, but big is necessary."

## (Alice Korngold; BRAC Is The Largest Global Anti-Poverty Organization, And It's A Secret; 5/15/2011; Fast Company)

**Appendix Twenty-One**

**Paris Declaration on Aid Effectiveness**

The Paris Declaration and Accra Agenda for Action are founded on five core principles, born out of decades of experience of what works for development, and what doesn't. These principles have gained support across the development community, changing aid prectice for the better:

It is now the norm for aid recipients to forge their own national development strategies with their parliaments and electorates (ownership); for donors to support these strategies (alignment) and work to streamline their efforts in-country (harmonisation); for development policies to be directed to achieveing clear goals and for progress towards these goals to be monitored (results); and for donors and recipients alike to be jointly responsible for achieveing these goals (mutual accountability). Beyond its principles on effective aid, the Paris Declaration (2005) lays out an action-oriented roadmap to improve the quality of aid and its impact on development. It puts in place a series of specific implementation measures and establishes a monitoring system to [assess progress](http://www.oecd.org/document/44/0,3746,en_2649_3236398_43385196_1_1_1_1,00.html) and ensure that donors and recipients hold each other accountable for their commitments. The Paris Declaration outlines the following five fundamental principles for making aid more effective:

**1. Ownership:** *Developing countries set their own strategies for poverty reduction, improve their institutions and tackle corruption.*

**2. Alignment:** *Donor countries align behind these objectives and use local systems*.

**3. Harmonisation:** *Donor countries coordinate, simplify procedures and share information to avoid duplication.*

**4. Results:** *Developing countries and donors shift focus to development results and results get measured.*

**5. Mutual** **accountability:** *Donors and partners are accountable for development results*

**Appendix Twenty-Two**

**World Health Organization Criteria for Social Determinants of Public Health**

World Conference on Social Determinants of Health Technical Paper

A range of specific responsibilities and tasks can be identified:

• Understanding the political agendas and administrative imperatives of other sectors

• Building the knowledge and evidence base of policy options and strategies

• Assessing comparative health consequences of options within the policy development

process

• Creating regular platforms for dialogue and problem solving with other sectors

• Evaluating the effectiveness of intersectoral work and integrated policy-making

• Building capacity through better mechanisms, resources, agency support and skilled and

dedicated staff

• Working with other arms of government to achieve their goals and in so doing advance

health and well-being

**Appendix Twenty-Three**

**Example of an Evaluation Protocol**

**Ad Council**

# Research & Evaluation Procedures

The central focus of the Ad Council’s mission is to make a measurable difference in society. Program evaluation is a critical component of every Ad Council campaign. In order to assess a campaign’s impact, we conduct research that encompasses a dashboard of indicators.

### Strategic Research & Evaluation

For every Ad Council campaign, extensive research is conducted during a campaign’s formative stage.  During this stage, our main goal is to garner insights that will inform an effective communications strategy.

**Literature review:** Before conducting any primary research, the team is briefed by the campaign sponsor on the issue.  We then conduct an extensive literature review, focusing on the nature and scope of the issue, other communications programs that have addressed this issue, and the most current empirical and public opinion studies.

**Exploratory research:** The pro bono advertising agency and the Ad Council conduct primary qualitative research among members of the target audience.  The research objectives are to determine how target audience members perceive and define the issue, and to identify who among the target audience is most likely to act on communications related to the issue.  These insights, along with the literature review, help the team to develop the campaign strategy.  We employ a variety of qualitative methodologies during the exploratory phase, which may include traditional focus groups or one-on-one interviews, in-home discussion groups, ethnographies, vlogging, or other methods.  We typically travel to at least 2-3 markets in a given project.  At times, we supplement the qualitative research with a survey, to test initial hypotheses or possible strategic directions among a quantitative sample.

**Communications checks:** After the campaign’s creative concepts have been developed and approved by the Ad Council, we go back to consumers in order to gauge their response to the advertising.  Here we seek to learn whether the work will be effective in gaining the target audience’s attention, conveying the main message, and persuading the audience to take action. This research can also help us determine the most effective concept among competing ideas.  The research is usually qualitative in nature, typically a series of focus groups in several markets.  Consumers are exposed to the advertising concepts in rough format (storyboards, scripts, etc.) and asked for their feedback.  On occasion, we also conduct quantitative copy tests of the advertising in finished format, where a larger panel of consumers are exposed to the ads and asked for their feedback.

**Peer Review:** Throughout this process, we receive feedback and approvals from two groups of experts: 1) the campaign sponsors, who are the issue experts; and 2) the Campaign Review Committee (CRC), who are the communications experts.  The CRC is a panel of top advertising executives assembled monthly by the Ad Council to review and approve campaign strategies, creative concepts, and final advertising materials.

### Post-launch Assessment

The Ad Council assesses a campaign’s ROI using a variety of measures.  Each campaign has a customized evaluation plan, agreed upon by all parties during the strategy stage.  Generally speaking, most campaigns employ the following types of metrics:   
  
**Media measurement:** Since we do not buy any media time or space, the Ad Council provides quarterly estimates of a campaign’s donated media support.  We employ several monitoring services to track media outlets’ support, including TV, radio, Web, newspaper, magazine, out-of-home and alternative media vehicles.  Based on monitoring service reports and self-reports from some media companies, we are able to estimate the number of ad placements donated to a particular campaign, the impressions generated, and the monetary value of each of these placements.  Also, we measure news media coverage generated by our public relations initiatives and events.   
  
**Consumer response:** Most Ad Council campaigns include a website and/or a toll-free number where consumers can get more information or become involved.  We monitor website traffic trends, calls to the toll-free number, brochure requests/downloads, email sign-ups, etc.  For campaigns that rely heavily on digital media we pay particular attention to sophisticated online metrics, including detailed website analytics, engagement with social media, and online buzz trends.   
  
**Tracking survey:** The Ad Council conducts a national benchmark survey of the target audience prior to the release of the campaign, followed by annual postwave surveys.  A tracking study gauges trends over time among the target audience.  Measures include awareness of the issue, recognition of the advertising, and relevant attitudes and behaviors.  While improvements in awareness and favorable attitudes are important, the gold standard for campaign assessment is usually a statistically significant increase in the desired behavior, as self-reported by the target audience.   For most surveys, we employ online methodology, recruiting respondents from large opt-in research panels.  For campaigns with audiences that are difficult to reach online, we employ telephone methodology.   
  
**Externally monitored data trends:** Government agencies and nonprofits occasionally conduct research that is related to an Ad Council campaign issue (e.g., childhood obesity rates, or high school dropout rates).  While we often cannot connect these large data trends directly back to an Ad Council campaign, monitoring these data provides a helpful context.

### Campaign Impact

In conjunction with our partner ad agencies and government/nonprofit sponsoring organizations, we conduct qualitative and quantitative research to guide the strategic and creative development of our campaigns. Campaign impact is evaluated via media exposure, consumer response, web analytics, and national pre- and post tracking surveys that measure awareness, attitudinal and behavioral shifts among the target audience.

**Ad Council campaigns assess return on investment** through several empirical measures:

* Donated media support (through media monitoring services and media company self-reporting);
* Awareness of campaign messages (measured through tracking surveys);
* Consumer response (direct response to our marketing efforts by visiting a campaign website or calling a toll-free number, as measured by website analytics and call center records). This may also include consumer activity on social networking sites (emailing spots, commenting on a campaign, participating in an online game or other activity);
* Word of mouth around a campaign issue (measured through tracking surveys and online buzz metrics);
* Shifts in relevant attitudes and/or behaviors (measured through tracking surveys, where attitudinal and behavioral scores are correlated with ad recognition scores);
* Empirical data collected by outside agencies (such as government data on drunk driving crashes).

(<http://www.adcouncil.org/default.aspx?id=480>; accessed 6/22/2011)

**Appendix Twenty-Four**

**Example of an excellent policy-oriented evaluation**

Kate Bird, David Booth, and Nicola Pratt; Food Security Crisis in Southern Africa: The Political Background to Policy Failure; August 2003; Forum for Food Security i n Southern Africa; <http://www.odi.org.uk/resources/download/1607.pdf>

**Appendix Twenty-Five**

**Sample Independence Governance Assessment Following a Crisis**

**ACORN, 12/7/2009, by Rick Cohen**

<http://www.nonprofitquarterly.org/index.php?view=article&catid=149%3Arick-cohen&id=1650%3Aan-independent-governance-assessment-of-acorn-the-path-to-meaningful-reform&format=pdf&option=com_content&Itemid=117>

**Appendix Twenty-Six**

**Questions for evaluators**

It may be useful for evaluators to refer to a checklist of questions commonly

asked in evaluations. A list of potential questions is presented below (drawing

heavily from Martin, L. ‘Planning and Organising Useful Evaluations’,

UNHCR, January 1998).

**Understanding the situation**

– How was the initial situation assessed? What access was there to the affected

populations? What data was collected, and how was it analysed?

– What would have happened if no assistance had been given?

– Were lives in danger?

– What are the security implications of providing/not providing assistance?

Were these considered?

– Are new movements of people/other significant events possible? Have

contingency plans been made for this?

– How have beneficiary groups been identified? Using what characterisation?

Were the special needs of certain groups (children, female-headed households,

disabled, ethic minorities) considered? Are any groups (unintentionally)

favoured over others as a consequence of the way the programme was set-

up? Are beneficiaries unfairly assisted in relationship to host/other

populations?

– Was the impact of the emergency and the humanitarian programme on host

populations and the local environment considered?

– What background research was carried out on the culture and history of

beneficiary groups? How were information gaps filled?

– What efforts were made to obtain reliable disaggregated population figures?

– What range of solutions were considered to the problem?

– Were other agencies included in analysis of the problem and the drawing up

of response plans? Were response plans ‘holistic’ or very much determined

by the mandate of individual organisations?

– What analysis has been undertaken of possible long-term solutions to the

problems? – What is the quality of this analysis?

– Were preparedness activities carried-out?

**Reviewing the plan**

– Have lessons been learnt from earlier programmes?

– Did original plans contemplate exit strategies?

– Were constraints to implementation adequately assessed from the outset?

– Were all critical factors considered, including the political, economic, and

cultural factors, as well as the weather and logistical considerations? What

subsequent changes took place in the institutional, political and economic

context? What impact did these changes have on programmes?

**Monitor assumptions**

– What major assumptions were made at early stages of the response? Did

these turn out to be well founded? If not, in what respect?

– Were there major considerations that should have been foreseen but were

not?

– Were the programme objectives sufficiently clear? Were alternative

approaches to achieving the objectives considered? Are objectives sufficiently

explicit so that their achievement can be assessed and impact determined?

– Are the activities within the mandate of the organisation concerned?

– Did activities correspond to actual needs? Are they consistent with solutions?

– What were the main inputs, outputs and activities?

– Do programme plans take into account longer-term (re-)integration of

beneficiaries?

– Do activities discourage initiative and lead to dependence?

– What was the administrative, personnel and logistic back-up like? Did short-

comings in these areas have any significant impact on performance?

– Was an assessment made of the links between assistance and protection

activities?

– Did staff have the experience and expertise to carry-out the activities

envisaged?

– Did headquarters and field staff agree on programme aims, priorities and

implementation strategies? Did lack of agreement cause any substantial

difficulties?

– Have cost considerations been a part of the planning and design process?

**Design**

– Does the operational/conceptual model/ logical framework underlying the

intervention make sense? Will the activities proposed achieve the desired

outputs? Will these outputs lead to the desired changes (impact)? Have

reasonable assumptions been made about what may affect the projected

outcome? Have important assumptions been left out or ignored?

**Involvement of beneficiaries**

– Were the interests of beneficiaries adequately taken into account in the design

and implementation of the programme? Or were procedures to canvass the

opinion of beneficiaries a token gesture, which over-represented the views

of the controlling elite, or those who were best able to articulate their views?

– Were beneficiary groups equitably and transparently identified for targeting

and were they involved in the negotiation of beneficiary status?

– Was the aid provided sufficient for need at the time and did it relate to

livelihood and coping strategies?

– Was aid culturally acceptable and appropriate?

– Did aid meet negotiated targets?

– Was aid delivered in a form that allowed ease of transport for the recipient?

– Was the delivery system transparent?

– Was the delivery of aid targeted directly to those in need, taking into account

limitations on their ability or inability to travel?

– Was there a rapid response by agencies to failures in the delivery system?

– Did aid allow people to make autonomous decisions about their own lives?

– Was human dignity maintained in agency dealings with beneficiaries?

– Did agencies actively seek and listen to the beneficiaries’ views on aspects

of the aid system?

– Were the security and safety of the beneficiaries high priorities?

**Assessing the context and linkages**

– What are the government’s objectives, and how do they differ?

– How can different agendas (whether between agencies, between agencies

and government or between government and other national forces) complicate

the response? Have measures been taken to mitigate the negative impact of

different agendas, or to try and bring those with different agendas together?

– Have national resources been used appropriately?

– Are donors in agreement about programme objectives?

– Was the best-placed organisation (in terms of knowledge of the country,

skill profile, capacity, etc.) used to carry out the programme?

– Are activities well coordinated? Are sectoral responsibilities successfully

divided between partners?

– Are plans well communicated (to other partners, governments and

beneficiaries)?

**Analysing Implementation**

– Were assumptions accurate?

– Were plans followed? If not, why not?

– Have objectives changed during implementation? Why? On the basis of

what analysis or what events?

– Are assistance activities contributing to protection?

– Are lines of authority and reporting clear?

– Are there staff or skill shortages? What has been done about these?

– Did staff receive adequate briefing?

– Did activities match needs?

– Did beneficiaries participate in programme implementation? What was the

impact on implementation of their (non-)participation?

– Have there been delays in implementation? How were these dealt with?

What was the impact of such delays?

– Has there been coordination and information-sharing with other agencies,

governments, donors, beneficiaries? Has this avoided duplication and led

to sensible sharing of tasks between actors?

– Was decision-making undertaken in a timely manner and delegated to the

appropriate level?

– Have correct financial and administrative procedures been followed?

– Are activities consistent with policies, recognised good practice and

guidelines?

– What problems have emerged, for what reasons and how have they been

dealt with?

– What have been the strengths and weaknesses of implementation?

– Where could activities have been more cost-effective?

– What training of partners has taken place? What capacity has been left

behind?

– Are programme activities regularly monitored and evaluated?

– Were measures to provide security and protection adequate?

– Were those most in need prioritised for assistance?

– Did activities contribute to solutions? In what way?

– Was social conflict prevented or limited?

– Will activities continue once agencies pull out?

– Were needs overlooked or neglected?

– What is the analysis of the current situation? How are events expected to

develop?

– Were commitments honoured? Did the assistance make a difference? What

were the most successful activities? And the least successful? Would

alternative forms of assistance been more effective?

– What were the long-term impacts of assistance?

– What will happen once agencies pull out?

– How do the achievements compare to plans and objectives?

– Were resources effectively and efficiently used? Do the results justify the

costs?

– Was assistance significant or of marginal value to beneficiaries?

– Were objectives relevant and worth pursuing?

– Where is improvement needed?

– What new programme, policies, activities, systems or guidelines are

required?

– What lessons can be learned?

**Appendix Twenty-Seven**

Exploring Humanitarian Law (EHL), International Committee of the Red Cross

<http://www.alnap.org/pool/vacancies/icrc-ehl-evaluation-tor(1).doc>

Scope of Evaluation

1. **Value for Operations**: How do delegations describe EHL's value and do they value the programme for its support to other ICRC objectives (i.e. positioning, cooperation, protection, etc.)?

* Is the "continuum" logic valid and useful (increase respect for the law and improve relations with communities affected by violence, by exposing people to IHL and Red Cross/Red Crescent action from their youth onwards)?
* Has EHL proved to be a useful tool for relationship-building with political authorities and/or with communities affected by violence?
* Can EHL form part of the ICRC's prevention response to identified humanitarian problems in context?

1. **Value for Interlocutors**: Is EHL effective and relevant to priorities and policies of education authorities and National Societies?

* Does EHL respond to an interest in and a demand for IHL/humanitarian education?
* What is EHL's added value in formal and non-formal education?
* Does EHL achieve the expected learning outcomes?
* Is EHL universal in its appeal? Can it be contextualized and remain pedagogically sound?

1. **Value for Cooperation**: How does EHL contribute to ICRC's relationships with National Societies?

* Building on the results of a prior external evaluation, which looked into the ICRC's investment in developing the capacity of National Societies to promote law and fundamental principles, why do National Societies consider EHL as an example of a successful joint endeavour?
* How does EHL contribute to delegations' cooperation objectives?
* How do National Societies understand ICRC's role in promoting and implementing EHL?
* Where/ how does EHL fit within National Societies' programmes?

1. **Defining Success**: Is sustainability a success factor and if so, what are the criteria for ensuring sustainability of EHL in a given context?

* In light of answers to the 3 other questions of the evaluation, is sustainability a success factor?
* What are the external conditions for sustainability?
* What are the internal conditions for sustainability?

Evaluation Methods

The evaluation will consist of desk research combined with field trips to 4 to 6 contexts, among the following: South eastern Europe (Bosnia and Herzegovina, Serbia, Macedonia), China, Iran, Peru, South Africa and the United States of America.

Contexts have been chosen in accordance with criteria that are relevant to the outcome of the evaluation: regional distribution, the level of implementation of the programme, the relative importance of the programme for the concerned ICRC region, the level of involvement of National Societies, and ICRC presence in the context.

The evaluation method should be further defined by the evaluator in their inception report. The evaluator is asked to devise rigorous assessment tools combining qualitative and quantitative data analyses. The following stages are required:

* desk research
* interviews with key ICRC staff involved with EHL implementation, at headquarters and in the field
* interviews with key external stakeholders including education authorities and National Society staff

If feasible and deemed desirable for the outcome of the evaluation, the evaluators ought to consider interviewing teacher-trainers, school administrators and students, through school visits.

**Appendix Twenty-Eight**

**Model Code of Ethics for Public Endeavors**

**Ethical Responsibility to Provide Service**

While reservations may exist regarding the provision of planning services, if consumers desire such services and if there is a reasonable expectation that interventions can be beneficial, then we have a professional obligation to address this demand.

**Do No Harm**

“In complex emergencies, to be effective (or at least to limit potential harm), relief workers must understand the underlying systems of production, social services, politics and markets because some relief interventions can unintentionally undermine self-sufficiency and increase vulnerability…This problem has plagued the relief community from its earliest days.” (Sue Lautze; Feinstein International Famine Center; Saving Lives

And Livelihoods: The Fundamentals Of A Livelihoods Strategy; Tufts University;

March 1997; See Cuny, F., (1983), Disasters and Development, Intertect Press, Dallas)

Any intervention needs to be considered within the context of the potential benefits to be accrued from it. Only by considering both the potential risks and the possible benefits can we appropriately evaluate a proposed intervention. The simple presence of risk does not necessarily preclude the use of an intervention if it is sufficiently justified by the potential benefits. With interventions that have a reasonable likelihood of being beneficial for the public, the important issue becomes for the planner to understand the nature of the risks, to minimize the risks to the extent possible, to fully inform members of the public as to the nature of the risks within the context of the possible benefits, and then to allow the people to make informed decisions about their intervention options.

**Provide Effective Interventions**

**Provide Equitable Interventions**

**Do justice**: ensure that the expected benefits are distributed equitably, and that burdens are not high, adverse, and disproportionate on any one group.

**Ensure diversity and inclusiveness**

Practice nondiscrimination.

**Be a responsible third party**

The third party, separate from decision-maker(s) or stakeholders, assists in designing and managing a process that pursues a variety of goals beneficial to the stakeholders and decision-maker(s). The third party is “a protector of the process.” (Moore, 1998) This means that she has a fiduciary role in defining, with substantive (if not total) stakeholder co-determination, the purpose, parameters, procedures and expectations for what will be mutually acceptable and beneficial procedures for engaging on an issue or pending government action…the core of third party is as a “helper of all” who guards against participants or interlopers eroding the value of the participation/ negotiation process (Carpenter, 1995). (John B. Stephens and Maureen Berner (2011) "Learning from Your Neighbor: The Value of Public Participation Evaluation For Public Policy Dispute Resolution," *Journal of Public Deliberation*: Vol. 7: No. 1, Article 10.   
http://services.bepress.com/jpd/vol7/iss1/art10)

**Don’t Intervene Beyond the Boundaries of Competence**

**Professional Accountability and the Opportunity for Redress of Grievances**

The targeted population should be clearly informed prior to beginning an intervention about the regulatory agencies and professional associations governing the planner’s work.

**Obtain Informed Consent**

**Consider the ability to intervene in a crisis**

**Report malfeasance, violations of law, threats of violence, and child and domestic abuse**

**Know your customer.**

An organizationmust be aware of whom it is doing business with in the areas of export control, anti-money laundering, and anti-terrorism.

**Establish professional and personal boundaries**

**No conflict of interest**

**Example:**

**Code of conduct for the International Red Cross and Red**

**Crescent Movement and NGOs in disaster relief**

1. The humanitarian imperative comes first

2. Aid is given regardless of the race, creed or nationality of the

recipients and without adverse distinction of any kind. Aid

priorities are calculated on the basis of need alone

3. Aid will not be used to further a particular political or religious

standpoint

4. We shall endeavour not to act as instruments of government

foreign policy

5. We shall respect culture and custom

6. We shall attempt to build disaster response on local capacities

7. Ways shall be found to involve programme beneficiaries in the

management of relief aid

8. Relief aid must strive to reduce future vulnerabilities to disaster

as well as meeting basic needs

9. We hold ourselves accountable to both those we seek to assist

and those from whom we accept resources

10. In our information, publicity and advertising activities, we shall

recognise disaster victims as dignified humans, not hopeless

objects

**Appendix Twenty-Nine**

The Community Development Resource Association in South Africa describes the “particularly undevelopmental global development industry” as characterized by:

* the need to urgently disburse money;
* a tendency to focus on product rather than process;
* higher respect for suits and ties than rages and bones;
* a proactive rather than responsive orientation;
* centralized and hierarchical decision-making; and
* bureaucratic and instrumental rigidities, practiced largely (if unconsciously) for the benefit of those who intervene.

**Appendix Thirty**

**Paris Declaration on Aid Effectiveness**

Signed by a great many countries and international organizations in 2005. The Paris Declaration broke new ground for achieving greater aid effectiveness on the basis of shared principles and measurable time-bound indicators.” (Stefan Meyer and Nils Sjard Schultz, *Paris to Accra: Building the Global Governance of Aid* (Madrid: Fride, August 2008), 16) It specifies indicators against which donor and beneficiary countries should measure their progress in achieving the five key Declaration Principles. It looks at the responsibility of developed and developing countries for delivering and managing aid. The phrase ‘Aid effectiveness’ refers to the efficacy of official development assistance (ODA) in achieving economic development in recipient nations. The document has been criticized because the Declaration is technically orientated, does not take into account the political dimensions of aid, and is designed to guide development assistance and not relief and stabilisation efforts. The Declaration has also led to a focus by the international community on the processes of managing aid rather than on the impact of aid. (Rebecca Roberts; Reflections on the Paris Declaration and Aid Effectiveness in Afghanistan; Discussion Paper; Afghanistan Research and Evaluation Unit; April 2009; http://reliefweb.int/sites/reliefweb.int/files/resources/A23234BD23762DFC4925759E001FFD8C-Full\_Report.pdf)

Ownership, Harmonisation, Alignment, Results and Mutual Accountability

Principles to increase the impact aid has in reducing poverty and inequality, increasing growth, building capacity and accelerating achievement of the Millennium Development Goals (MDGs). These include:

Scale up for more effective aid

Adapt and apply to differing country situations

Specify indicators, timetable and targets

Monitor and evaluate implementation

Partner countries exercise effective leadership over their development policies, and strategies and co-ordinate development actions

Donors base their overall support on partner countries’ national development strategies, institutions and procedures

Donors align with partners’ strategies

Donors use strengthened country systems

Conduct diagnostic reviews.

Partner countries strengthen development capacity with support from donors

Strengthen public financial management capacity

Strengthen national procurement systems

Untie aid: getting better value for money

Donors’ actions are more harmonised, transparent and collectively effective

Donors implement common arrangements and simplify procedures

Reduce the number of separate, duplicative, missions to the field and diagnostic reviews; and promote joint training to share lessons learnt and build a community of practice.

Complementarity: more effective division of labour

Excessive fragmentation of aid at global, country or sector level impairs aid effectiveness. A pragmatic approach to the division of labour and burden sharing increases complementarity and can reduce transaction costs.

Incentives for collaborative behaviour

Reform procedures and strengthen incentives—including for recruitment, appraisal and training—for management and staff to work towards harmonisation, alignment and results.

Delivering effective aid in fragile states

Adapt to environments of weak ownership and capacity and to immediate needs for

basic service delivery.

Promoting a harmonised approach to environmental assessments

Managing resources and improving decision-making for results

Managing and implementing aid in a way that focuses on the desired results and uses information to improve decision-making.

Mutual Accountability

Donors and partners are accountable for development results

Enhance mutual accountability and transparency in the use of development resources.

Reinforce participatory approaches by systematically involving a broad range of development partners when formulating and assessing progress in implementing national development strategies.

Provide timely, transparent and comprehensive information on aid flows so as to enable partner authorities to present comprehensive budget reports to their legislatures and citizens.

Aid flows are aligned on national priorities.

Avoid parallel implementation structures

The Busan meeting on aid effectiveness added these:

Ownership   of   development   priorities   by   developing   countries.

Inclusion  of  new  actors  on  the  basis  of  shared  principles  and  differential commitments

Civil  society  organisations  (CSOs)  play  a  vital  role  in  enabling  people  to  claim their  rights,  in   promoting   rights-­‐based   approaches,   in   shaping   development   policies   and   partnerships,   and   in  overseeing  their  implementation.

Improve  the  coherence  of  our  policies  on  multilateral  institutions,  global  funds  and

programmes.

Address  the  problem  of  insufficient  delegation  of  authority  to  field  staff.

Promoting  sustainable  development  in  situations  of  conflict  and  fragility

Prioritise  legitimate  politics,  people’s security, justice, economic foundations and revenues and fair services.

Partnering  to  strengthen  resilience  and  reduce  vulnerability  in  the  face  of  adversity

Invest   in   shock   resistant   infrastructure   and   social   protection   systems   for   at-­‐risk communities.  In  addition,  we  will  increase  the  resources,  planning  and  skills  for  disaster  management  at  the  national  and  regional  levels.

Support   the   implementation   of   institutional   and   policy   changes   led   by   developing  countries,   resulting   in   effective   resource   mobilisation   and   service   delivery,   including  national  and  sub-­‐national  institutions,  regional  organisations, parliaments  and  civil  society.

Support  the  development  of  improved  evidence  on  institutional  performance  to inform policy  formulation,  implementation  and  accountability,  led  by  developing  countries.

Engage  with  representative  business  associations,  trade  unions  and  others  to improve  the legal,   regulatory   and   administrative   environment   for   the   development   of   private investment;  and  also  to  ensure  a  sound  policy  and regulatory  environment  for  private sector  development,  increased  foreign  direct investment,  public-­‐private  partnerships,  the strengthening  of  value  chains  in  an equitable  manner  and  giving  particular  consideration  to national  and  regional dimensions,  and  the  scaling  up  of  efforts  in  support  of  development goals.

Enable   the   participation   of   the   private   sector   in   the   design   and   implementation   of development  policies  and  strategies  to  foster  sustainable  growth  and  poverty  reduction.

Further  develop  innovative  financial  mechanisms  to  mobilise  private  finance  for shared development  goals.

Promote “aid for trade” as an engine of sustainable development, focusing on outcomes and  impact,  to  build  productive  capacities,  help  address  market  failures,  strengthen access  yo  capital  markets  and  to  promote  approaches  that  mitigate  risk  faced  by  private  sector  actors.

Combating  corruption  and  illicit  flows. Implement  fully  our  respective  commitments to  eradicate  corruption,  enforcing  our  laws and  promoting  a  culture  of  zero tolerance  for  all  corrupt  practices.  This  includes  efforts  to  improve   fiscal   transparency,   strengthen   independent   enforcement   mechanisms,   and extend protection  for  whistleblowers.  Strengthening   anti-money   laundering   measures,   addressing   tax   evasion,   and   strengthening   national   and international   policies,   legal   frameworks   and   institutional   arrangements   for   the   tracing, freezing   and   recovery   of   illegal   assets.   This   includes   ensuring   enactment   and

implementation  of  laws  and  practices  that  facilitate  effective  international cooperation.

Climate  change  finance. Support  national  climate  change  policy  and  planning  as  an  integral  part  of  developing countries’ overall national development plans.

<http://www.oecd.org/dataoecd/15/3/46874580.pdf>

**Appendix Thirty-One**

**Fundamental Principles of International Charity**

Developed by the Treasury Guidelines Working Group of Charitable Sector

Organizations and Advisors

March 2005

http://www.usig.org/PDFs/Principles\_Final.pdf

1. Consistent with the privilege inherent in their tax-exempt status, charitable organizations must exclusively pursue the charitable purposes for which they were organized and chartered.

2. Charitable organizations must comply with both U.S. laws applicable to charities and the relevant laws of the foreign jurisdictions in which they engage in charitable work.

Charitable organizations, however, are non-governmental entities that are not agents for

enforcement of U.S. or foreign laws or the policies reflected in them.

3. Charitable organizations may choose to adopt practices in addition to those required by law that, in their judgment, provide additional confidence that all assets—whether resources or services—are used exclusively for charitable purposes.

4. The responsibility for observance of relevant laws and adoption and implementation of

practices consistent with the principles contained herein ultimately lies with the governing board of each individual charitable organization. The board of directors of each charitable organization must oversee implementation of the governance practices to be followed by the organization.

5. Fiscal responsibility is fundamental to international charitable work. Therefore, an organization’s commitment to the charitable use of its assets must be reflected at every level of the organization.

6. When supplying charitable resources, fiscal responsibility on the part of the provider

generally involves:

a. in advance of payment, determining that the potential recipient of monetary or in-kind

contributions has the ability to both accomplish the charitable purpose of the grant and

protect the resources from diversion to non-charitable purposes;

b. reducing the terms of the grant to a written agreement signed by both the charitable

resource provider and the recipient;

c. engaging in ongoing monitoring of the recipient and of activities under the grant; and

d. seeking correction of any misuse of resources on the part of the recipient.

7. When supplying charitable services, fiscal responsibility on the part of a provider involves taking appropriate measures to reduce the risk that its assets would be used for non-charitable purposes. Given the range of services in which organizations engage, the specific measures necessarily vary depending on the type of services and the exigencies of the surrounding circumstances. The key to fiscal responsibility, however, is having sufficient financial controls in place to trace funds between receipt by the service provider and delivery of the service.

8. Each charitable organization must safeguard its relationship with the communities it serves in order to deliver effective programs. This relationship is founded on local understanding and acceptance of the independence of the charitable organization. If this foundation is shaken, the organization’s ability to be of assistance and the safety of those delivering assistance is at serious risk.

**Appendix Thirty-Two**

**20 Tips On Evaluation**

The Beverly Foundation

http://beverlyfoundation.org/wp-content/uploads/20-Tips\_on\_Evaluation.pdf

In evaluating a volunteer driver program it is important to...

1. Review evaluation criteria.

2. Design appropriate evaluation methods.

3. Involve staff, drivers and riders in the evaluation process.

4. Use an outside evaluator if possible.

5. Analyze driver feedback information.

6. Analyze rider feedback information.

7. Analyze staff reports.

8. Analyze operation data.

9. Analyze ride data.

10. Analyze cost data.

11. Analyze performance data.

12. Review reaction to and support for the program.

13. Review funding community support for the program.

14. Review feedback from volunteer organizations.

15. Review appropriateness of risk management strategy.

16. Review effectiveness of driver recruitment.

17. Review effectiveness of rider recruitment.

18. Measure impact of the program on volunteer drivers.

19. Measure impact of the program on riders.

20. Analyze data to determine program efficiency, effectiveness, and impact.

**Appendix Thirty-Three**

**Measures of Effectiveness**

In presenting current US Army doctrine on ‘Support Operations’, Army Field Manual 3-0

states that commanders should establish quantitative measures to evaluate mission

effectiveness:

‘10-43. With supported agencies and governments, commanders establish **measures of**

**effectiveness to gauge mission accomplishment**. Measures of effectiveness focus on the

condition and activity of those being supported. Those that are discrete, measurable, and link cause and effect help commanders understand and measure progress and success. In famine relief, for example, it may be tempting to measure effectiveness by the gross amount of food delivered. In some cases, this may be an acceptable gauge. However, a better one may be the total nourishment delivered, as measured by the total number of calories delivered per person per day, or the rate of decline of deaths directly attributable to starvation. Measures of effectiveness depend on the situation and require readjustment as the situation and guidance change.’

( Department of the Army (2001) ‘Support Operations’, Chapter 10 in Field Manual (FM) 3-0, Operations, Washington, DC: Headquarters, Department of the Army, June)

**Suggested measures of effectiveness for evaluation of PRTs**

**Activity Suggested measures of effectiveness (MoEs) Suggested indicators/data sources**

Security

Changes in humanitarian security

Attacks on humanitarian workers

(Source: ANSO)

Humanitarian access

Changes in overall security environment

Swisspeace FAST Reports

Inter-militia intervention rate

Percentage of known inter-militia disputes in which PRT intervened

Inter-militia mediation success rate

Percentage of inter-militia disputes (in which PRT intervened) that were successfully mediated/resolved

Reduction in levels of lawlessness/banditry Reported incidents of banditry

Changes in land area under poppy cultivation in a specific region

Land area under cultivation [hectares]

People’s ambient security expectations (including in returnee areas)

Surveys of public opinion on perceptions of security (should not be undertaken by military)

Number of ANP personnel/trainees that can be supported by PRT

Number of ANP personnel

Security (‘hearts and minds’ activities)

Increased acceptance of military’s mission

Views expressed by local community members of military’s mission and

role (interviews should not be undertaken by military)

Improved co-operation between military and local population

Number of engagements by local community representatives with military in liaison capacity

Force protection Security intelligence on PRTs provided directly by members of the

local community

Reconstruction Support to road network construction

Km road/year (with reference to any government targets)

Use of roads

Local employment in PRT-funded reconstruction projects

Number of workers (full/part-time)

Employee salaries compared to local salaries

Component of reconstruction needs (per province) addressed by PRT activities

Comparison of needs (from ‘Securing Afghanistan’s Future’) and PRT

activities

Construction of facilities directly or under contract from PRT to support deployment of ANP (eg, police stations)

New/refurbished facilities (possibly measured by square metre) constructed by PRT in province

Support to central government

Relative authority of central

government in PRT catchment area

Qualitative data: who collects ‘taxes’/customs duties?

Who provides security?

Physical infrastructure available to support

Offices are staffed and functioning

Support by PRT to local government councils Number of functioning local

government bodies supported by PRT

Relief Operations

Relief delivered in specific sectors, according to verified need (eg, health; food and

nutrition; water and sanitation; and education sectors)

Sphere standards and indicators

(Gerard McHugh and Lola Gostelow; Provincial Reconstruction Teams and Humanitarian–Military Relations in Afghanistan; 2004; Save the Children; <http://www.cdint.org/documents/SCUK_PRT_Paper_FINAL.pdf>)

**Appendix Thirty-Four**

**Gender in Evaluations**

“Has gender received substantial attention in the planning?

Superficial references to gender do not necessarily vouch for gender mainstreaming.

Useful documents to consider:

Toolkit on mainstreaming gender equality in EC development cooperation

http://ec.europa.eu/europeaid/sp/gender-toolkit/en/content/toolkit.htm

**Mainstreaming gender equality through the project approach:**

Highlighting gender equality during implementation phase:

- Assess whether changes in the political, legal, economic or social context of the project have affected the assumptions about gender roles and relationships made at the beginning of the project. Phenomena such as migration, civil unrest or the HIV/AIDS pandemic can all affect existing roles and relationships and may require adjustments to the project.

- Assess whether project results are being delivered in a gender-equitable manner as was originally planned, and make adjustments if that is not the case.

- Assess whether information on project achievements/results is disaggregated by sex, and whether action is taken to redress inequalities and shortfalls

- Assess whether capacity is being built within the project structure and among stakeholders to ensure that gender equality achievements can be maintained after the project ends.

- The Task Manager will ensure that skills in gender analysis are included as a component in all TOR, and that the Gender Mainstreamed logframe is used as a management tool.

OECD Gender Policy Marker [www.oecd.org/dataoecd/4/23/39903666.pdf](http://www.oecd.org/dataoecd/4/23/39903666.pdf)

**OECD Gender Policy Marker:**

“An activity should be classified as gender equality focused (score Principal (2) or Significant (1)) if it is intended to advance gender equality and women’s empowerment or reduce discrimination and inequalities based on sex.”

A “principal” mark is given if the gender equality is the crucial, defining objective of a project.

A “significant” mark can be attributed to a project, which in its pursuit of another primary objective, integrates gender considerations as a substantial factor into the project design.

- EU Roadmap for equality between women and men (2006-2010)

http://europa.eu/scadplus/leg/en/cha/c10404.htm

- March 8th 2007 EC Communication on Gender Equality & Women Empowerment in Development Cooperation:

COM(2007) 100 final

http://eur-lex.europa.eu/LexUriServ/site/en/com/2007/com2007\_0100en01.pdf

(European Commission, EuropeAid Co-operation Office; Result Oriented Monitoring Handbook; July 2009; <http://ec.europa.eu/europeaid/how/ensure-aid-effectiveness/documents/rom_handbook2009_en.pdf>)

**Appendix Thirty-Five**

**Equity Evaluation Sample**

RACE Matters

ORGANIZATIONAL SELF-ASSESSMENT

Annie E. Casey Foundation

http://www.aecf.org/upload/publicationfiles/organization\_self\_assessment.pdf

“Why Should I Use This Tool?

Because unequal opportunities and racial inequity are deeply embedded

and usually not intended, producing equitable opportunities, operations

and results requires being intentional.

What Will the Tool Accomplish?

It raises organizational awareness, starts focused conversations,

contributes to the development of equity action plans, and tracks

organizational change. Organizations that care about these issues

can produce early results by using this tool.

How Do I Use It?

Answer each question by circling the response that most closely

applies. (For the section on staff competencies, decide if your focus

is your own unit or the entire organization.)

\_ Add up the numbers associated with each answer to get your Racial

Equity Score.

\_ Use the chart at the end of the tool to find out what your Racial

Equity Score means for your next steps.

STAFF COMPETENCIES

1. Staff are trained in and are knowledgeable at the 101 level about the range of barriers to equal opportunity and the depth of embedded racial inequities—how they are produced and how they can be reduced.

0=None 1=Some 2=Almost All 3=All

2. Staff have a deep level of understanding about barriers to opportunity and embedded racial inequities in their special area of focus—including critical data and information about how inequities are produced and how they can be reduced.

0=None 1=Some 2=Almost All 3=All

3. Staff are comfortable and competent about discussing issues of barriers to opportunity and embedded racial inequities with relevant individuals and groups.

0=Rarely 1=Sometimes 2=Almost Always 3=Always

4. Staff exhibit cultural competence in interactions with diverse groups.

0=None 1=Some 2=Almost All 3=All

5. Staff disaggregate data by race in all analyses.

0=Rarely 1=Sometimes 2=Almost Always 3=Always

6. A racial equity analysis is applied to policy issues.

0=Rarely 1=Sometimes 2=Almost Always 3=Always

7. A racial equity analysis is applied to practice issues.

0=Rarely 1=Sometimes 2=Almost Always 3=Always

8. Written materials reflect a knowledge and understanding of barriers to opportunity

and embedded racial inequities.

0=None 1=Some 2=Almost All 3=All

9. Staff can articulate the costs of failing to address barriers to opportunity and embedded

racial inequities.

0=None 1=Some 2=Almost All 3=All

ORGANIZATIONAL OPERATIONS

1. Removing barriers to opportunity and disparity/disproportionality reduction are explicit goals of the work and are articulated in a mission/vision statement.

0=No 1=Moving In That Direction 2=Yes

2. The unit has an internal team that guides the ongoing work of removing barriers to

opportunity and reducing racial disparity/disproportionality.

0=No 1=Moving In That Direction 2=Yes

3. The organization’s goals of reducing barriers to opportunity and racial disparities/

disproportionality are reflected in resource allocations.

0=No 1=Moving In That Direction 2=Yes

4. Investments promote capacity-building and asset-building for people and communities

of color.

0=Rarely 1=Sometimes 2=Almost Always 3=Always

5. Results of investments show opportunity for all and a reduction in racial disparities/

disproportionality.

0=Rarely 1=Sometimes 2=Almost Always 3=Always

6. The organization has a deliberate plan to develop and promote the leadership of staff of color.

0=No 1=Moving In That Direction 2=Yes

7. The organization has regular trainings and discussions at the staff and/or board levels about removing barriers to opportunity and reducing racial disparities and disproportionality, both internally and externally.

0=No 1=Moving In That Direction 2=Yes

8. The organization regularly assesses workforce composition by race/ethnicity and

develops/implements strategies for increasing diversity at all levels.

0=No 1=Moving In That Direction 2=Yes

9. The environment of the organization (food, art, holiday activities, etc.) is multicultural.

0=No 1=Moving In That Direction 2=Yes

10. The organization has a mechanism in place to address complaints about barriers to

opportunity and racial inequities in the workplace.

0=No 1=Moving In That Direction 2=Yes

NOW: Add up the numbers associated with each answer to get your Racial Equity Score.

My Racial Equity Score is \_\_\_\_\_\_\_\_.

FINALLY: Use the chart below to see what your next steps should be.

RACIAL EQUITY SCORE

<20

20–29

30–39

40–49

NEXT STEPS

Become intentional

Make an emphasis on racially equitable results explicit in your unit’s/organization’s

mission statement, and evaluate performance with this emphasis as a criterion.

Build staff/organizational capacity

If fewer points are in the area of Staff Competencies, identify opportunities for

staff to better understand embedded racial inequities – how they are produced and

maintained, and how they can be eliminated. If fewer points are in the area of organizational operations, identify policies and procedures that should be improved to promote racially equitable results.

Fine-tune staff/organizational capacity

See which items are scored lowest, and work on them.

Mentor others!

Because one unit’s or organization’s success in promoting opportunity for all and reducing disparities is likely to be tied to others’ performance, use what you’ve learned to

help advance a racial equity approach for critical partners.”

**Appendix Thirty-Six**

**US Government Evaluation of a Funding Application**

“EDA will use criteria e. through g., below, to evaluate the

extent to which the application satisfies the criteria.

a. The proposed project’s feasibility and the applicant’s organizational

capacity and ability to successfully execute the project and achieve desired

outcomes;

b. The integration of resiliency, as defined in section I.C. of this FFO, into

the project scope of work;

c. How the project will be sustainable and durable;

d. The project’s alignment with a relevant strategic economic development

or disaster recovery plan;

e. The likelihood that the project will foster job creation and promote private

investment in the regional economy;

f. The strength of the nexus of the project’s scope of work to the major

disaster designation; and

g. The extent to which the project is responsive to the needs of communities

impacted by disasters declared in FY 2011 regarding disaster relief,

long-term economic recovery, and restoration and enhancement of

infrastructure.”

(FY 2012 Disaster Relief Opportunity; Initial announcement of Federal Funding Opportunity; March 28, 2012; <http://www.eda.gov/pdf/FY2012_Disaster_Relief_Opportunity_FFO_FINAL.pdf>)

**Appendix Thirty-Seven**

**Evaluation of Scaling Up**

There are many unanswered questions about the trade-offs between scale and comprehensive, locally driven solutions. A few come to mind:

Are local areas better served when a number of organizations come in from the outside with their scalable solutions?  (one for education, one for water, one for entrepreneurship, one for health?)

Are efficiencies in service delivery increased or decreased at a community level by one model or the other?

Does the introduction of new scalable models into a community bring new positive ways of working?

What type of organization tends to hire more locals?

What kind of an organization does a better job at building community capacity ––within the eco-system?

What model or combination of models is best for the future in terms of sustained positive change for clients and community development?

(Kriss Deiglmeier;  
Executive Director, Center for Social Innovation; Stanford University; What Do We Lose with Scale? 7/05/2012; <http://csi.gsb.stanford.edu/what-do-we-lose-scale>)

**Appendix Thirty-eight**

**Evaluating access to justice**

Measuring Access to Justice project.

Central tenets are that the theoretical framework of measuring accessibility of justice will include three attributes of processes that are accessed. These are costs (out of pocket expenses, but also the costs of time spent, costs of delay, and emotional costs), outcome quality and procedural quality.

Measuring Access to Justice is a research project co-funded by the Hague Institute for the Internationalization of the Law, University of Tilburg and University of Utrecht aiming to develop, operationalize and pilot test a framework for measuring access to justice in international and in national settings. It is carried out by a research group of scholars from University of Tilburg and University of Utrecht.

M. Barendrecht, P. Kamminga and J. H. Verdonschot, Indications of Legal Needs: A

Bottom Up Approach to Rule of Law and Access to Justice, (2007).

M. Barendrecht, J. Mulder and I. Giesen, How to Measure the Price and Quality of

Access to Justice, (2006).

**Appendix Thirty-Nine**

**Example of Indicators**

**Individuals with Disabilities in Education Act**

U.S. Department of Education

[Indicators are an evaluative tool]

**List of Indicators**

[Indicator 1: Graduation Rates](http://nichcy.org/laws/idea/partb/indicators-partb/indicator1)  
Percent of youth with individualized education programs (IEPs) graduating from high school with a regular diploma compared to percent of all youth in the State graduating with a regular diploma.

[Indicator 2: Drop out Rates](http://nichcy.org/laws/idea/partb/indicators-partb/indicator2)  
Percent of youth with IEPs dropping out of high school compared to the percent of all youth in the State dropping out of high school.

[Indicator 3: Participation and Performance on Statewide Assessments](http://nichcy.org/laws/idea/partb/indicators-partb/indicator3)  
Participation and performance of children with disabilities on statewide assessments.

[Indicator 4: Suspensions and Expulsions](http://nichcy.org/laws/idea/partb/indicators-partb/indicator4)  
Rates of suspension and expulsion

[Indicator 5: Participation/Time in General Education Settings](http://nichcy.org/laws/idea/partb/indicators-partb/indicator5) (LRE)  
Percent of children with IEPs aged 6 through 21:

A.   Removed from regular class less than 21% of the day;  
B.     Removed from regular class greater than 60% of the day; or  
C.     Served in public or private separate schools, residential placements, or homebound or hospital placements.

[Indicator 6: Preschool Children in General Education Settings](http://nichcy.org/laws/idea/partb/indicators-partb/indicator6) (Pre-School LRE)  
Percent of preschool children with IEPs who received special education and related services in settings with typically developing peers (e.g., early childhood settings, home, and part-time early childhood/part-time early childhood special education settings).

[Indicator 7: Preschool Children with Improved Outcomes](http://nichcy.org/laws/idea/partb/indicators-partb/indicator7)   
Percent of preschool children with IEPs who demonstrate improved:

A.    Positive social-emotional skills (including social relationships);  
B.    Acquisition and use of knowledge and skills (including early  language/communication and early literacy); and  
C.   Use of appropriate behaviors to meet their needs.

[Indicator 8: Parental Involvement](http://nichcy.org/laws/idea/partb/indicators-partb/indicator8)  
Percent of parents with a child receiving special education services who report that schools facilitated parent involvement as a means of improving services and results for children with disabilities.

[Indicator 9: Disproportionate Representation in Special Education that is the Result of Inappropriate Identification](http://nichcy.org/laws/idea/partb/indicators-partb/indicator9)  
Percent of districts with disproportionate representation of racial and ethnic groups in special education and related services that is the result of inappropriate identification.

[Indicator 10: Disproportionate Representation in Specific Disability Categories](http://nichcy.org/laws/idea/partb/indicators-partb/indicator10)  
Percent of districts with disproportionate representation of racial and ethnic groups in specific disability categories that is the result of inappropriate identification.

[Indicator 11: Timeframe Between Evaluation and Identification](http://nichcy.org/laws/idea/partb/indicators-partb/indicator11) (Child Find)  
Percent of children with parental consent to evaluate who were evaluated and eligibility determined within 60 days (or State established timeframe).

[Indicator 12: Transition Between Part C and Part B](http://nichcy.org/laws/idea/partb/indicators-partb/indicator12)  
Percent of children referred by Part C prior to age 3 and who are found eligible for Part B who have an IEP developed and implemented by their third birthdays.

[Indicator 13: Post School Transition Goals in IEP](http://nichcy.org/laws/idea/partb/indicators-partb/indicator13) | Percent of youth aged 16 and above with an IEP that includes coordinated, measurable, annual IEP goals and transition services that will reasonably enable the child to meet the post-secondary goals.

[Indicator 14: Participation in Postsecondary Settings One Year After Graduation](http://nichcy.org/laws/idea/partb/indicators-partb/indicator14) | Percent of youth who had IEPs, are no longer in secondary school and who have been competitively employed, enrolled in some type of postsecondary school, or both, within one year of leaving high school.

[Indicator 15: Timely Correction of Noncompliance](http://nichcy.org/laws/idea/partb/indicators-partb/indicator15)  
General supervision system (including monitoring, complaints, hearings, etc.) identifies and corrects noncompliance as soon as possible but in no case later than one year from identification.

[Indicator 16: Resolution of Written Complaints](http://nichcy.org/laws/idea/partb/indicators-partb/indicators16)  
Percent of signed written complaints with reports issued that were resolved within the 60-day timeline, including a timeline extended for exceptional circumstances with respect to a particular complaint.

[Indicator 17: Due Process Timelines](http://nichcy.org/laws/idea/partb/indicators-partb/indicator17)  
Percent of fully adjudicated due process hearing requests that were fully adjudicated within the 45-day timeline or a timeline that is properly extended by the hearing officer at the request of either party.

[Indicator 18: Hearing Requests Resolved by Resolution Sessions](http://nichcy.org/laws/idea/partb/indicators-partb/indicator18)  
Percent of hearing requests that went to resolution sessions that were resolved through resolution session settlement agreements.

[Indicator 19: Mediations Resulting in Mediation Agreements](http://nichcy.org/laws/idea/partb/indicators-partb/indicator19)  
Percent of mediations held that resulted in mediation agreements.

[Indicator 20: Timeliness and Accuracy of State Reported Data](http://nichcy.org/laws/idea/partb/indicators-partb/indicator20)  
State reported data (618 and State Performance Plan and Annual Performance Report) are timely and accurate.

**Appendix Forty**

**Examples of Evaluation Failures**

1. BIAS ANYONE? The same party that implements the intervention (recognition, training, or any other) also evaluates its effectiveness..  
  
2. OVERSELLING NUMBERS. The results support a small correlation, impacting less than 5% of the variability … and yet the evaluators sell the results as groundbreaking wonders.  
  
3. HISTORY LESSONS. The evaluators fail to isolate the results of the intervention from other events (for instance, someone marvels at the great results of a sales program and ignores positive changes in the economy or the launch of a new product).  
  
4. SINGLE CHILD. The evaluators fail to establish a control group and only report on improvements experienced by the treatment group (i.e., the group experiencing the training, recognition program, etc.).  
  
5. OUT OF CONTROL. Alternatively, the evaluators establish a control group but design the research study so poorly that the results are meaningless.   
  
Dr. Cris Wildermuth, SPHR  
Community Chair, Linked:HR

**Appendix Forty-One**

|  |
| --- |
| **Criteria for Evaluation Measures** |
| **Relevant:** Provide direct feedback on the outcomes of local implementation. |
| **Feasible:** Leverage use of credible, commonly collected data for cost effective reporting that is flexible enough to be implemented in various local contexts. |
| **Timely:** Capture actions and outcomes that are currently relevant, while looking forward to the future. |
| **Useful:** Provide significant value to help cities and counties make decisions and address local priorities. |
| **Systemic:** Draw attention to the preferred future, while offering a metric that measures true progress toward that achievement goal; and, where possible, will satisfy performance reporting for multiple goals to highlight the integrated nature of sustainability. |
| **Reliable:** Provide a consistent reflection of achievement or performance across communities regardless of community characteristics, facilitating comparisons between communities. |
| **Valid:** Represent the concepts and underlying phenomena that are embodied in the STAR Objective accurately. |
| Adapted from “Indicators and Information Systems for Sustainable Development: A Report to the Balaton Group” by Donella Meadows, 1998  **Appendix Forty-Two**  **Financial, procedural, and policy constraints** Progress of comprehensive everglades restoration plan evaluated ***Date:***  June 27, 2014  ***Source:***  National Academy of Sciences  ***Summary:***  Although planning for Everglades restoration projects has advanced considerably over the past two years, financial, procedural, and policy constraints have impeded project implementation, says a new congressionally mandated report.  140627113058-large  This map shows locations of CERP and CERP-related restoration and pilot projects.  Credit: National Academy of Sciences  [[Click to enlarge image]](javascript:void(0))  Although planning for Everglades restoration projects has advanced considerably over the past two years, financial, procedural, and policy constraints have impeded project implementation, says a new congressionally mandated report from the National Research Council. Timely authorization, adequate funding levels, and creative policy and implementation strategies are needed to achieve restoration benefits and to expedite implementation of the Central Everglades Planning Project. Climate change and the invasion of nonnative plant and animal species further challenge the Everglades system. The impacts of climate change -- especially sea-level rise -- provide a stimulus to accelerate restoration efforts, the report adds.  The report is the fifth in a series of biennial evaluations of the Comprehensive Everglades Restoration Plan (CERP), a multibillion dollar project launched in 2000 with the goal of reversing the ecosystem's decline and creating a water system that simultaneously serves the natural, urban, and agricultural needs of southern Florida.  According to the report, restoration progress to date has been modest and focused along the edges of the ecosystem. The Central Everglades Planning Project, initiated in October 2011, recommends a suite of projects that would provide substantial new water flow to the central Everglades, equivalent to approximately two-thirds of the new water envisioned in CERP. The committee that wrote the report urged CERP planners and policymakers to find solutions to expedite the project's implementation in order to avert further degradation of the ecosystem's core. Without such solutions, water redistribution may not be feasible until 2035 or later, and with the envisioned funding level of $100 million per year, construction would not be completed for approximately four decades.  The report found that the infrequency of Water Resources Development Acts (the congressional mechanism for authorizing CERP projects exceeding $25 million), the availability of funding, and cost-sharing challenges have impeded CERP progress over the past two years. The Water Resources and Reform Development Act of 2014 -- the first authorization in seven years -- enabled four additional CERP projects to proceed with federal funding, although the Central Everglades Planning Project was not completed in time to be included in the legislation.  The Integrated Delivery Schedule, which lays out construction plans for the next decade, needs to be revisited to incorporate the newly authorized projects and the Central Everglades Planning Project with existing restoration efforts, the report says. Given limited funding, all projects cannot be advanced equally, and planners should consider factors such as possible climate change and sea-level rise to determine which projects have the greatest potential for restoration benefits.  Sea-level rise has already increased saltwater intrusion into Everglades freshwater habitats and urban water supplies, and potential future changes in temperature and precipitation may affect the timing, volume, and quality of freshwater and the distribution of species, as well as increase agricultural water demands. Although they pose a challenge to restoration efforts, climate change and sea-level rise are reasons to accelerate restoration to enhance the ecosystem's ability to adapt to future changes. For example, improvements in water depths could promote the accumulation of peat in Everglades wetlands, reducing coastal wetland loss caused by sea-level rise.  The report recommends that climate change be incorporated into adaptive management planning at both the project scale and in systemwide goals, and that planners build flexibility into the design so new knowledge and improved climate change projections can be incorporated as they become available and future restoration efforts can be adjusted appropriately.  The report also states that planners should consider the implications of restoration activities on nonnative species. Invasive plants and animals displace native species and disrupt ecosystem structure and function, and some projects may affect the extent and abundance of nonnative species. The committee found that although there has been good coordination of invasive species management at the project level, strategic coordination over management and research priorities is lacking.  Setting effective priorities for managing invasive species requires a comprehensive understanding of all nonnative species present in the Everglades and their impacts and threats, as well as those of impending or possible new arrivals. Research is needed to determine which species could reasonably be predicted to have considerable ecological impacts. A strategic early detection and rapid response system that addresses all areas, habitats, and species is needed, the report states.  Scientific research provides the knowledge and tools that can help decision makers ensure that the resources invested in Everglades restoration are used wisely. Long-term monitoring collects useful information for understanding how projects are changing ecosystem conditions. A comprehensive re-evaluation of restoration-related monitoring is needed to determine its adequacy considering budget pressures, extended implementation time frames, and potential impacts of climate change and sea-level rise. In addition, renewed attention to science coordination and communication is needed, which includes adequate funding, staff, and a clear charge to address scientific priorities for restoration.  Report: <http://www.nap.edu/catalog.php?record_id=18809>  **Story Source:**  The above story is based on [materials](http://www8.nationalacademies.org/onpinews/newsitem.aspx?RecordID=18809" \t "_blank) provided by **[National Academy of Sciences](http://www.nas.edu" \t "_blank)**. Note: Materials may be edited for content and length.  **Cite This Page**:   * [MLA](http://www.sciencedaily.com/releases/2014/06/140627113058.htm?utm_source=feedburner&utm_medium=email&utm_campaign=Feed%3A+sciencedaily%2Ftop_news+%28ScienceDaily%3A+Top+News%29#citation_mla) * [APA](http://www.sciencedaily.com/releases/2014/06/140627113058.htm?utm_source=feedburner&utm_medium=email&utm_campaign=Feed%3A+sciencedaily%2Ftop_news+%28ScienceDaily%3A+Top+News%29#citation_apa) * [Chicago](http://www.sciencedaily.com/releases/2014/06/140627113058.htm?utm_source=feedburner&utm_medium=email&utm_campaign=Feed%3A+sciencedaily%2Ftop_news+%28ScienceDaily%3A+Top+News%29#citation_chicago)   National Academy of Sciences. "Progress of comprehensive everglades restoration plan evaluated." ScienceDaily. ScienceDaily, 27 June 2014. <www.sciencedaily.com/releases/2014/06/140627113058.htm>. |