

1. DFID Evaluation Policy- Meta analysis and Synthesis Studies
2. Beneficiary Counting
3. Citizens Report Card and Exit Polls
4. Mapping exercise of Civil Society Organisations
5. How to estimate sample size for control group
6. What is Black Box in M&E Field
7. Question on Framing Evaluation Questions
8. Pre and Post Test for Measuring Training Effectiveness
9. M&E News Summary

1. DFID Evaluation Policy- Meta analysis and Synthesis Studies

Reviews of the findings from (and approaches used in) existing DFID evaluations would be a useful work stream - noting however, that such evaluations are drawn on when an agency performs a sector or country-level review.

In the context of synthetic reviews I would like to draw your attention to (1) the Campbell Collaboration colloquium in May which will focus on development interventions - the call for proposals closes on January 15 (<http://www.campbellcollaboration.org/>); and (2) 3ie's support for synthetic reviews. 3ie will shortly be announcing its first awards to conduct reviews, and putting out another call for proposals during January.

Howard White

~

I would distinguish between meta-evaluation, and synthesis studies/reviews, of past evaluations: In my view a meta-evaluation is an evaluation of the evaluation processes that have been used, whereas a review or synthesis study would be more concerned with analysing and aggregating judgements about the findings from a set of evaluations. The latter is what the Campbell Collaboration (and others like it) are all about, as far as I understand. I am not recommending more of these, useful as they are.

My recommendation was for a meta-evaluation. If you plan to decentralise evaluations, and then provide support for this process, then there must be some form of theory-of-change (tacit or explicit) about what is then supposed to happen e.g. about the subsequent quantity, quality, relevance, ownership, etc of the evaluations that are decentralised. As far as I know DFID EvD have not yet carried out this type of evaluation, yet decentralised evaluations are a major part of their evaluation strategy.

Rick Davies

~

While undertaking the study "Mapping of NGO Monitoring and Evaluation Practices" (2008) among Danish Development NGO for Danida, I found that evaluation of evaluation practices have generally not been much explored among governmental and non-governmental actors alike, nor are there many guiding frameworks in this field.

The study may be downloaded from the website of the Ministry of Foreign Affairs, Denmark (www.um.dk) and search for title.

Hanne Lund Madsen

~

I need help regarding a meta-analysis study that I have commissioned. I am in charge of evaluating a big programme where I work. Many provinces and departments that report to this programme have done their own evaluation studies on the programme.

I then want to commission a study that will consolidate all these reports and inform me what they

have found, how many of them were of good quality, how many of them are poor. Of those that are of good quality, what were their findings about the programme?

What kind of study is appropriate? Is it a meta-analysis or meta-review? Is there a difference? I will appreciate any help and any material that I can read on the subject.

KG

~

"How many of them were of good quality, how many of them are poor. Of those that are of good quality, what were their findings about the programme?"

Meta analysis is a method of statistically combining the results of good studies. I'm not sure it is a method of telling which studies are the good ones. Meta-review seems to be a review of reviews (see below).

However, it kind of sounds like you are looking for researchers who are very good, who can review the evaluations and tell you which are good, not good, and what the good ones say. So I don't know that you need a 'study' but rather a review.

Gene Shackman

~

There are three concepts related to the question re meta-analysis. These include meta-evaluation, meta-analysis and evaluation synthesis.

Meta-evaluation (sometimes called pooling evaluation) is the evaluation of the quality of evaluations using various standard criteria. Some times it is also known as 'the evaluation of evaluations'. D.Stufflebeam (1974) defined meta-evaluation as a procedure for describing an evaluation activity and judging it against a set of ideals concerning what constitutes good evaluation. In other words, meta-evaluation means assessing the merit of evaluation efforts. Similarly, Michael Q. Patton (1997: 143) defined meta-evaluation as a process to evaluate a given evaluation studies based on the profession's standards and principles. The motive for conducting meta-evaluation in Patton's terms is to ensure an independent and credible review of an evaluation's strengths and weaknesses. In this meta-evaluation context we are then dealing with questions such as: was the evaluation well done? Is it worth using? Did the evaluation meet professional standards and principles? and so forth (see also Patton, 1997: 193, 333).

Meta-analysis is a statistical way of aggregation of the results of individual studies. Glass et al (Glass, 1976) identified three characteristics of meta-analysis. First, meta-analysis is quantitative and uses numbers and statistical techniques for organizing and extracting valuable information that is nearly incomprehensible by other methods. Second, meta-analysis does not tend to evaluate the quality of existing studies. However, meta-analysis attempts to record various aspects of research methodologies for the existing studies to identify their relationship to study findings. Third, meta-analysis aims to compare existing studies and to seek general conclusions across studies.

An evaluation synthesis is a "content synthesis of various evaluation studies (largely qualitative)", i.e. a global evaluation/cross-section analysis (Widmer, 2003). An evaluation synthesis requires a thematic clustering among the evaluation studies upon which it is based, but has a less narrow focus due to its qualitative approach.

It is important to note that mostly meta-evaluation precedes meta-analysis. This is because evaluations should be included in to a meta-analysis only when they fulfill certain level of minimum quality standards. Hence, based on your explanation you are in need of conducting meta-evaluation (to evaluate which evaluations are good/poor) and meta-analysis (if the evaluations are quantitative) or evaluation synthesis (if the evaluations are largely qualitative).

Hailemichael Taye

2. Beneficiary Counting

I am interested in examples or experiences with beneficiary counting for multi-sector programming. I recently started a new M&E position with an international organization with programs in 186 countries ranging from emergency relief to livelihoods. Consequently, the scale and scope of programming and resultant beneficiaries is tremendous, ranging from those who receive a house to condoms, or indirectly benefitting from a public lifeguard, establishment of a disaster early warning system, or construction of a hospital. I am reminded of a caution from one of my early M&E mentors of the "McDonald's approach," which measures the billions served but says nothing about the quality (nutritional value) of what is being served. However, there is strong interest in being able to collect beneficiary data and report on collective outreach. Therefore, I welcome any input on beneficiary counting for multi-sector programming.

Scott Chaplowe

~

Interesting observation "measures the billions served but says nothing about the quality". Let me ask even a more basic question. Leave aside quality, **what does normal practice actually count?** Different methods can give different counts. E.g. If an individual opts for 25 days in a Cash for Work program, the individual can be counted once through one system and 25 in yet another system, generating wide variations in count projections. Unless and until methods for beneficiary counts are fully disclosed, such statistics touted end up as being a little disingenuous.

Rajan Alexander

~

Don't know if this answers your question but am focusing more on the qualitative. depending on the sampling size, i wonder if there is a benchmarking conducted prior to the start of the intervention. If it is already in existing, one use participatory M & E, using MSC to capture qualitative impact or even emerging impact of the project. What can be captured may be direct or indirect - I remembered, reviewing a PRA report, on a forest management project. However, as i have reviewed other documents attached (health reports and education report), apparently, there was also the issue of toxicity, as evidenced in the yellowish skin of the children.

~

In my discussion with the community leaders, i said, based on their problem tree, this will be another project impact - other than the usual measurement (e. number of trees, decrease in water turbidity, change behaviour in environment conservation, etc). While there might be mandatory indicators, i think using the MSC or even AI, other project impact that is demand driven based on a specific situation will be captured.

Ann Lily

First off, I want to thank all for the time in your responses. The points raised are extremely relevant, and they reflect the challenges of "beneficiary" counting. I do like StC's approach of using "reach," which kind of disengages the term from impact/quality. The question was asked, why count beneficiaries. At this stage, the interest is in reporting on collective outreach. After instilling a culture of collecting and reporting on such output counts, the vision is to scale up to higher level, outcome indicators (KAP -knowledge, attitudes, practice). Double counting is certainly a challenge, that we are well-aware of from the multiple program setting of the tsunami recovery effort, where there was much overlap among programs and "beneficiaries." Of course, this is complicated by counting direct beneficiaries, i.e. those who have received a shelter, well, or septic system, versus indirect beneficiaries, i.e. those benefitting from the establishment of an early warning system, pre-positioned emergency stocks, or the construction of a hospital. At this stage, I agree that an important part of the solution is to be clear and consistent on the guidelines for conducting such counts. Again, thanks to all for the valuable input.

Scott Chaplowe

Further to the discussion on the topic Beneficiary count, I would like to share some of our experiences in this regard; Beneficiary count is an important global activity of *Save the Children* and we annually count the "Reach" covered by our main 4 Result sectors. We also believe that it this is the count of coverage and has nothing to do with impacts or quality of program, for which we adopt different methodologies. I can not go into the details of how we count, however, I

would like to highlight some key issues which we encountered and developed some remedial measures:

- 1) **Clear concepts and guidelines** who to count, who are direct and who are indirect beneficiaries are very important. These may vary from program to program, and even activities to activities, so this requires clear guidelines according to the program sector and types and nature of beneficiaries. At the top of it, we also need to know why do we need to count? If needed, then to what level of details and how these data will be used? This will of course depend on Agency priority, importance and need of Beneficiary data.
- 2) **Keeping good data base of beneficiaries** reached by the Program sectors is important to correctly count the beneficiaries. Unless you have some kinds of database/templates and people are regularly keeping it update, beneficiary count is a difficult task
- 3). **Duplication issues:** Different types of duplication issues exist: duplication created by repeated beneficiaries of same activity, duplication of same people but by different activities, overlapping of multiple programs, duplication between direct and indirect counting and so on. . No need to say that if you are not careful, you will be counting the same beneficiaries so many times. We have found that most difficult one is to eliminate cross-sectoral duplications or duplication created by multiple programs in the same geographic areas. Good knowledge of programs, and nature of beneficiaries, good records of beneficiaries data and agreed methods to eradicate duplication are necessary to avoid the duplication. But note that when we have multi-sector programs, we can only estimate the duplication in many cases, there is no accurate methods as yet. The important precaution you need to take is that , as mentioned below, we do not cross beyond the population in our counting.
- 4) **Concepts and guideline orientation** not only to the M&E staff but all program staff who in fact deal with the beneficiaries is most important. Save the children has a global guideline on Beneficiary reach with concepts and sector-specific guides on how and who to count. If some of you are interested , I can forward.

Dinesh Uprety

~

You write: "I am reminded of coverage boasted by the Marshall Islands Health Department nutrition programs. They counted the number of attendees at their nutrition classes (presumably from attendance records) which exceeded the population of the country." That should not come as a shocker to most of us. Let say the class strength is 20 and the program is for 30 days. Assume 100% attendance every day. Assume further daily attendance per class is considered the basis for counting, not number of students passing the course. So for 20 students, the counting could be as high as 600! It gets even better if these figures are used for further "analysis". Total cost of training lets assume as \$6000. For the count of 600, average cost is a mere \$10 per individual (great "efficiency"!) while for a count of 20, it becomes as high as \$300 (poor "efficiency") I am not sure whether such distortions arise from pure ignorance or deliberate disingenuity practiced. What's sure is that this is sadly very much part of development reporting!

Rajan Alexander

3. Citizens Report Card and Exit Polls

My current task is setting up an M and E systems for a governance programme in Sierra Leone. Our focus is to link organisation working to create the space, and build meaningful communication, between citizens, civil society organization and state. The programme log frame and M and E Plan has been finalised and agreed by our donors. There are a number of tools I hope to develop for effective and routine monitoring of our interventions.

I am writing to seek information from members of the group to share the following tools and training manual with me:

1. **Citizens Report Card** – For evaluating of citizens satisfaction with basic service at the community level. We are looking at Access, Utilization and Satisfaction with service at the community level. Also we want to evaluate the relationships and interaction between local councils and their communities.

2. **Exit Polls** – For evaluating dialogue sessions, training and solidarity events
Thank you very much for your support.

Patrick Elogima Robin

~

Can any one in the group help me out with citizen report card and exit polls evaluation techniques. I am currently working on a governance programme and will want to evaluate citizens satisfaction with the quality of services and their interaction with local councils. We also want to evaluate solidarity events and training programmes.

Patrick Elogima Robin

~

You may also want to take a look at the Affiliated Network on Social Accountability (ANSA) Africa website (<http://www.ansa-africa.net>) for resources on social accountability and to link to a large network of social accountability initiatives and actors. The secretariat of this network has recently moved from the Human Sciences Research Council (<http://www.hsrc.ac.za>), a large public entity in South Africa to IDASA (<http://www.idasa.org.za>), which is a strong South African NGO involved in governance related matters. Both organizations have experience with Social Accountability methods.

Marlene Roefs

~

An e-mail based discussion group called Solution Exchange is supported by the United Nations in India. A Consolidated Reply on this discussion group is a good resource for developing Human Development/ governance Report Cards by citizens. I am giving below the web link to this document. Please contact me if you are not able to access this document for any reason:
<http://www.solutionexchange-un.net.in/decn/cr/cr-se-decn-07120701.pdf>

Pankaj

~

Am in the field right now conducting a mid-term review for one of our HIV/Aids projects. I am currently using a modem and the download/upload speeds are pretty slow. I will send you the design template and report of a CRC survey that I was one of the lead persons for on Govt service delivery in Kenya on Monday when I get back to the office on Monday. I hope that will not be too late.

Lewis Aritho

4. Mapping exercise of Civil Society Organisations

I need help in writing a brief concept of about 11/2 pages on how I view a mapping exercise of all CSOs operating in the field of HIV and AIDS in Swaziland which I am applying for as a regional enumerator. This mapping exercise is part of an Internal Organizational Assessment commissioned by UNAIDS through the Coordinating Assembly of Non-Governmental Organizations in Swaziland (CANGO).

Please help me out on the structure of this brief and what I write about. The deadline is 4th February 2009.

Sandile Ginindza

~

You might want to look at Net-Map (<http://netmap.wordpress>) as a tool for mapping your civil society organisations. It's a low tech method for mapping actors, formal and informal links, goals or actors and their influence on the issue at stake. Have a look at the website for a manual and case-studies that might help you develop your own strategy.

Eva Schiffer

~

I realize that you have a tight time frame, so I have not waited for a response from you on the number of projects involved. I have done what I am describing below with up to 30 projects, but it could be done with more.

Below is an example of how you can map multiple projects onto an outcomes model if that is what you want to do. You build the outcomes model according to the 13 Tips here <http://easyoutcomes.org/files/13tipsa4-205.pdf>

And then you can map your projects onto the common outcomes model. (you could map them onto various common models you can develop, including HIV related outcomes, outcomes related to the nationwide system dealing with HIV, specific population groups etc.)

A mock-up of what this looks like is available at <http://www.outcomesmodels.org/models/educationsector39.html>

Just click through all of the pages (use the green down arrow at the top of the page) and you will find an example of how the projects can be mapped onto the model.

And you can see here how you can put in the number of projects that map onto each outcomes to assist strategic discussions in the example

below <http://www.outcomesmodels.org/models/communityprojectmapping29.html>

These examples are web page models which have been created from within DoView outcomes and evaluation software. The first example is mocked up in a soon to be released no cost update version, which will allow images to be included.

The second example is in the current version of DoView. You can download a trial version of DoView from <http://www.doview.com/download.html> and you can download the second example by clicking on Download the DoView file of this model on the bottom of the screen when you are looking at the second example. You can get other models to play around with at <http://www.outcomesmodels.org>.

You can get a full workbook on how to do all this

from http://www.easyoutcomes.org/files/easyoutcomesworkbookv1-9sv_2008-4-11.pdf

[Disclosure: I have been involved in the development of DoView software]

Paul Duignan, PhD

~

I forgot to mention, that there is an example of a NGO network mapping exercise available online, here: [Davies 2007 "Research and Advocacy Organisations and their relations with issue-base coalitions in Ghana"](#) report on an online survey

at <http://www.mande.co.uk/ghana/Online%20survey%20of%20coalitions.doc>

Rick Davies

5. How to estimate sample size for control group

Dear All,

I joined this group two days ago and am pleased to contribute and share my skills.

There are various methods of calculating your sample size. However there are issues that you need to have in mind, these include Level of precision commonly referred to as sampling error, the level of confidence and the degree of variability or what you would call distribution of your study attributes in the population. Remember the more heterogeneous is your population then a larger size would be necessary to obtain the given precision.

Purity Njagi

Monitoring and evaluation Officer

On Tue, Feb 3, 2009 at 7:03 AM, jalandhar pradhan

<jpp_pradhan@yahoo.co.uk> wrote:

> Dear All,

> We are planning to conduct a baseline survey for an intervention project.

> The study is based on case-control design.

> So, could you please suggest me, is there any methodology to estimate the

- > number of sample/ village to be covered for control area for comparison
- > purpose.
- > Infact for project area we are estimating the sample size using the 'P1' and
- > 'P2' value.
- > Can we apply the same formula to estimate the sample size for control area?
- > Waiting for your valuable suggestions.
- > Regards,
- > JAlandhar

There are lots of methods for sample size calculation, but it depends on how sophisticated you want this to be. Least sophisticated is to assume fixed treatment effects throughout, in which case you can just use the usual sample size calculations to determine the overall N needed, and then allocate shares of the sample over the clusters (in this case, villages) in a manner that either optimizes precision or streamlines analysis by minimizing the use of weights (there's a tradeoff). In this case, overall sample size can be computed using standard sample size calculation software, such as the sampsi function in Stata. Precision or "streamlining" as above requires application of some basic sample design principles.

But if you think there may be some heterogeneity in treatment effects or that there may be random effects associated with clusters, then you would want to use some form of cluster design framework to determine optimal number of clusters and optimal number of units to be sampled within each cluster. There is a program called "Optimal Design" that performs these calculations:

<http://sitemaker.umich.edu/group-based/home>

Or, you can set up a custom simulation-based program to determine sample size. I've done this for some of my evaluation work. If you are interested, we can talk details.

Cyrus Samii
 Political Science
 Columbia University
[cgs81@columbia.edu](mailto:cds81@columbia.edu)

Burundi Survey: www.columbia.edu/~cgs81/burundisurvey/
 ISERP Statistical Consulting:
www.iserp.columbia.edu/services/statistical_consulting.html
 Comparative Political Economy Blog: cpecolumbia.blogspot.com

Greetings from snowy Connecticut--

Sample size calculations are the same, regardless of whether it is for the treatment or control group. Sample size is the function of the probability of the occurrence of the event, and the desired precision of the estimate. If you need to make comparisons between 2 groups, you are going to want relatively precise estimates otherwise it will be difficult to show that the treatment had a statistically significant impact on the population of interest.

Beth

6. What is Black Box in M&E Field

Dear All,

I am Imtiaz and working as M&E Officer. I am not clear of the term Black Box in the M&E field. Please guide and if you could furnish with some literature.

Imtiaz

~

Dear Hailemichael Taye

I hope you may have some clarification by reading the beautiful documents available in the weblinks.

http://www.wca-infonet.org/servlet/BinaryDownloaderServlet?filename=1018271895315_SJebellie_Paper8.pdf
<http://www.itrc.org/papers/benchmarkingirrigation/benchmarkingirrigation.pdf>

S.Srinivasan

~

Hi,

In M&E, I think black box is the area where actions are taken to design M&E system until and unless outputs, outcome, effects, impacts and/or results are not obtained. If we suppose a water reservoir, it usually has an inlet and then water goes through some process and after that it flows out through outlet. This process, which starts soon after the input and before it flows out of outlet, is a black box. In the project management cycle it carries the same concept behind. So it could be then the planning, implementation strategy, monitoring and evaluation, retreat and policy reform in a project before we get the required result of a project. Similarly, it could be the implementation strategy of M&E such as what tools to be used in M&E, how the tools will be used, who will use it, when will it be used, what data/information we need and from whom, at what size, at what time, what will happen to data, how and who will process it, what will be reported and to whom it will be and how can the report help the project. This is the black box in M&E where stakeholders such as designers, project staff and decision makers need to coordinate and participatorily put together efficient and effective approach and methodology to get the required result from M&E.

Colleagues can correct me if I am wrong 😊

Dad

~

To add on, I have cut and paste from World Bank nutrition resources. I remembered in one of the projects I was involved our team had a lingering concern on project evaluation focusing on anthropometric data collection alone during mid-term and end evaluation... Read below. In Vietnam, community health and nutrition workers observed that, despite comparable socioeconomic status, some children were growing adequately while most suffered from varying degrees of under-nutrition. Using the positive deviance* approach, the behavioural characteristics of "successful" households were identified. These behaviours included the collecting of shrimp and crabs while working in the fields and supplementing a child's diet with them. An educational campaign promoting such positive deviant behaviours was implemented. After several months, an evaluation was conducted to investigate the effect of the project on nutritional status. Evaluators were pleased to discover that nutritional status had substantially improved since the baseline. However, the evaluation collected only anthropometric data and neglected to determine whether the adoption of positive deviant behaviours had, in fact, taken place. While the overall improved growth of children in the project area is cause for celebration, the conclusion that this resulted from the educational campaign may have been incorrect. In fact, deworming medication had been introduced into this area during the same period and may have contributed substantially to the decreased rates of under-nutrition. Simply examining impact relegates the reasons for change to a "black box". Unfortunately, this approach to evaluation is all too common: improvement in impact indicators, where it is observed, is assumed to be attributable to the project without examining process (what we later will define as output and outcome) indicators. *Positive deviance refers to situations in which individuals or households are doing better than would be expected given their social and/or economic circumstances. Their time and resource allocation strategies may be worth disseminating more broadly.

Ann Lily Marie O. Uvero

~

The term "black box" came from the so-called "black box evaluation model." It is the origin of interventions model that develop under program theory. The model is usually presented as follows.

Inputs -> "Black box" -> Outputs

The black box could be anything, it can be an organization or an intervention. It is assumed that if one give inputs to it, it will produce outputs. In brief, black box is something that will "change"

inputs into outputs. In M&E, practitioners often equal black box with processes or activities. Also, it is considered that black box model is contextual in nature in which external environment (e.g. social, political, economy, environment, etc.) could affect inputs, outputs and the black box itself. It should be emphasised here that this is the primitive model of evaluation.

From this concept, then the logic model as follows developed afterward:

Inputs -> Intervention: Process/Activity -> Outputs -> Outcomes -> Impacts

I suggest you to read Weisz (1996) classic book of evaluation to learn more on this black box model, or other text on program theory.

Ari

~

My understanding, and use, of the term is that a "black box" approach is one which reports on outputs or outcomes without any analysis of the underlying causal mechanisms (program theory, logic etc) by which the program inputs have resulted in affecting those outputs/outcomes. Lest there be any doubt, this is a bad thing!

So in the area of impact evaluation, approaches relying solely on regression-based estimates from the coefficient on the project dummy variables, or control group-based estimates simply looking at mean differences in outcomes between treatment and control, are often guilty of adopting a black box approach. The draw back of this approach is that it misses possible policy lessons that come from unpacking the causal chain.

As an example, I was involved in the evaluation of a nutrition project in Bangladesh which found little impact on nutritional status, though allowing for impact heterogeneity (looking at how impact varies across groups) found that there was an impact on the most malnourished. The most malnourished were of course the target group, but the data showed substantial mis-targeting, with many of those who should benefit not doing so (Type I error), whilst many in the program should not have been there (Type II error). Why was this? Admission to the program was through growth monitoring, in which community workers weighed each child monthly and mapped the child's progress on a growth chart. It turned out that these community workers could not correctly read the growth charts to identify which children ought be admitted to the program. Hence a clear reason for low impact (children who will not benefit are getting the intervention), and a policy implication (improve the quality or training of community workers). (There were other reasons, I have picked one for illustrative purposes).

There is a brief discussion of black boxes in my pamphlet on impact evaluation:

<http://ideas.repec.org/p/pramprapa/1111.html> and the Bangladesh project is written up in

<http://ideas.repec.org/a/wly/jintdv/v19y2007i5p627-652.html> and

<http://ideas.repec.org/p/wpa/wuwpdc/0510004.html>

Howard White

7. Question on Framing Evaluation Questions

I just wanted to get your insight on the best way for framing evaluation questions. Is it appropriate to base the summative evaluation (and take evaluation questions) directly upon the project/program objective statements and the scope of work as defined in the contract/proposal/any official documentation?

If so, then theoretically what type of evaluation is that? (e.g. evaluation serving accountability only? or ...)

Look forward to your comments,

Sana

~

Allow me to ask you about purpose of evaluation before my cement. Do you want evaluation for project process, output evaluation or impact evaluation? Usually Project proposal should have its M & E plan and on the basis of that one can plan evaluation. Evaluation is usually done for process/output/result or may be impact. It depends upon researcher if he/she wants to evaluation process, output, result or impact. Some time objectives statement are too widened or ambitious and its difficult to frame out on objectives which may take more time to achieve in contrary to project timeframe.

Javed Pasha

~

The intended user of the evaluation i am referring to is the management of the center, running the project and the donor agency. The intended use of the evaluation findings is to learn about the project outcomes as a test-case for scaling up on similar initiatives (designing similar initiatives as pointed out by Mr. Williams) and from the donor point of view it is to serve accountability purpose. The center's management is interested on a kind of output evaluation and to some degree impact aswell, but really the focus is to see what has been achieved in terms of the objectives that were set. Therefore in such a case I would request suggestions that is it alright to start by taking the project objectives (both general and specific objectives and the planned scope of work, but define more specific sub-questions) for evaluating if the project has effectively achieved the set goals or not.

Sana

8. Pre and Post Test for Measuring Training Effectiveness

I hope you can help me on this issue. I am planning to conduct trainings, and thinking to use pre and post test as one of the tool to measure the effectiveness of the trainings. However recently I read in a Monitoring and Evaluation manual, it is not recommended to use pre and post test as the tool to measure the training effectiveness, I'm not quite understand why, since there is no argument to support this statement. Maybe there are any of you that have experience in using pre and post test to measure the trainings and can explain to me why it is not recommended. I also would like to know if there is another tools I can use to measure the training effectiveness.

Laura Ginting

~

One of the reasons behind why Pre-Post Test design are not advisable to be used for trainign effectiveness because of "intervention time". Changes that may take place because of training delivery would be extremely artificial at that time of measure. And it will lead you to conclude that the training is effective and we knew it is NOT. For Pre-Post design, intervention time is an element of primary importance. However, at some stance, the Feedback of the Training would be more appropriate which also depends on your evaluation instruments.

Becky P

~

Question - when you say you want to measure the effectiveness of training, what exactly do you want to measure? Whether, for example, people learned something from the training, or whether they took the training and did something with it. In either case, pre and post test seems the best way to determine any effects of training, along with all the other issues of random assignment, self selection, etc. that you also have to deal with.

Gene

~

Relating to your question, As far as I know, Pre and Post Test is used to measure the training outputs, that is to know whether the training participants had gain the knowledge given in the trainings. So, I guess we usually applied the tools in the first day of training (for pre test) and at the last day of the training (for Post test) - *am I right? Please do correct me if I'm wrong.* When I read the statement that says "The test is not recommended to be used for measuring the trainings" I assumed the person stated that is not recommending the M&E officers to use it for measuring the training outputs, and it makes me to ask Why? But from your opinion and other colleagues opinion, I can make conclusion that Pre and Post Test is still the best tool for measuring the training outputs (is the trainee has gained the knowledge given from trainings), and it still is the best as long as we do not use it for evaluation (for outcome, and impact). - *Please also correct me if I make the wrong conclusion.*

Laura Ginting

~

Yes, "Pre and Post Test is still the best tool for measuring the training outputs (has the trainee has gained the knowledge given from trainings)" At least at one particular time.
As Becky also wrote "Changes that may take place because of training delivery would be extremely artificial at that time of measure." That is, right after the training ended, people who do better appear to have learned the material, or have learned how to respond to way you want them to respond.

However, whether this learning translates to doing anything with it, after the training, that is another question, and you won't get an answer to that question by giving a survey right before training and right after training. You would need another survey or some kind of feedback later on, after people return to their 'home' settings, and had a chance to apply what they 'learned'.

Gene Shackman

~

Pre post approach can judge the delivery of training and NOT its effectiveness. The effectiveness of the training can be measured after the training is applied by the training recipients/ or the effect that the training bring to the organization (depending on the type of training).

Shaukat

~

I want to support what Shaukat has said. You can use pre and post test to measure training delivery and what new information the participants have acquired in the course of the training.

However the effectiveness of the training itself will take time. This will best be measure after sometimes, you can then measure if the training as produced the desired improvement in the delivery of the participants work with an eye on the objective of the training.

World Bank has a whole document on measuring the effectiveness of their training. the document is titled "Using training to build the capacity for development" you can download one or request for a copy.

Olanipekun Oluwasola G.(Nehi)

~

It is really an interesting discussion. However, all said and done, I would still go with pre and post evaluation. But we must be very clear what indicators we want to measure. For example if you want to measure the gender attitude among the trainees, you may have a pre training evaluation and post training evaluation and if you want to see the sustainability of the training, you may still want to to another follow up evaluation after certain period.

In any case, pre and post test evaluation is the most effective methods to measure effectiveness, either in change of behavior or in attitude, but we should be very clear what we really want to measure.

Ajay K Singh

But then my question would be that is it something measurable or evaluated "near" to the project end or soon after it. How do we see such kind of effectiveness if the project evaluation has to be done within one month of the project end (that includes the report writing time also).

Sana

~

Much as it may be too early to ascertain effectiveness of a training just after a few months, as M&E practitioners, through our monitoring I think we need to focus a little more on results of activities such as trainings. This is what Results Based Management (RBM) is all about. Instead of just looking at outputs of trainings, we can seek to establish through monitoring whether actually these trainings are having the desired results. By doing this, I think we can later on (during evaluation) be able to attribute certain changes in the project environment to our activities (such as trainings).

Titus G Tumusiime

~

There is a better way of assessing training sessions and measuring knowledge acquired (output) by trainees in the course of participation in training. Though, it's still pre-post test but this is administered on each day of training over presentations for that day.

Methodology

M&E officer will ask for at least five questions from each facilitator each day which will be administered on the participants before the training and after the training daily. On each day, previous daily performance should be displayed in a chart to show increase/decrease in knowledge on each sessions presented. The approach is actually tasking but will help to measure the following:

- Facilitators' skill in disseminating topic s/he planned to.
- Participants' knowledge gained/lost.
- Its makes participants actively participate in training sessions knowing that at the end of each day there will be test.

We should also note that if there is difficulty in measuring immediate result (output) of inputs, measuring outcome may be impossible. Therefore, we cannot conclude that "Changes that may take place because of training delivery would be extremely artificial at that time of measure," as Becky wrote.

We should also bear in mind that •MONITORING• is a continuous activity, after training people for a particular activity we should device means (tool) for ensuring that learning translate to doing something with it.

AWOLEYE JOSHUA OLATUNJI

~

The achievement of results is gradual i.e from input-output-outcome-impact. It is possible to see the outcome of your efforts within the lifetime of the projects which starts building up after achievement of the outputs within certain assumptions. That is if you are able to see the immediate results given certain assumptions then there is a likelihood of seeing the intermediate results and the outcome within the lifetime of the project. The duration of the project also matters and the kind of intervention being undertaken but it is possible to see outcomes being generated slowly by slowly as time elapses. Just think a bout process of making wine. The wine matures with time.

Alex

~

When it comes to evaluate training outcome, I guess there is no particular ideal time for it, since in my opinion the more participants feel the benefit from the training they've got (the given training is exactly what the participants' need), the behavior will change faster means the expected outcome will happen faster, and it is possible to be achieved within the project lifetime.

The problem is, sometimes we give training that is not needed by participants or they did not realize yet that the trainings will actually benefit them maybe in the future. This will cause the behavior changed slowly, means the expected outcome will also slowly to be achieved, sometimes took years. In this case, I have experience to make evaluation of the trainings at least in 6 months even though the outcome had not achieved yet, but the process towards achieving the outcome sometimes can be identified. Besides, usually donors would required to have evaluation report in every 6 months.

Perhaps others can discuss this too...

Laura

~

Activities (like trainings) can be measured monthly. If we talk about training outputs (ie. knowledge gained from training), I don't think we can measure only at post-test, as knowledge tends to decline over time. When I had worked in Vietnam, we did post-tests, but then re-tested a sample of training participants every 6 months to see if there was still good understanding. Often it was not, which forced us to re-design activities (often by doing refresher trainings, improved trainer-quality, reducing #s of training participants, etc.). Since our community development work was done in 15 years chunks, with impact evaluations done every 5 years, we measure outcomes (ie. changes in behavior) every 5 years. However, I came to the conclusion that this was not sufficient and that some outcomes could be measured annually. We did not want to wait 5 years to see if knowledge actually translated into behavior.

Greg

~

I think Javed has rightly added word Practice to ASK complementing the purpose of training but how effective could be training in term of Practice, Knowledge, Attitude and skill are totally different things to be measured in context of training effectiveness. How effective the training was delivered is something else depends upon its contents, methodology and audience understanding or its acceptance. Training effectiveness in term of practice or attitude can't be judged in pre and post training evaluation . It can't be easy to measure effectiveness of training in attitudinal change or practice of people right after completion of training as it would take time to have attitudinal change

Mr. Javed Iqbal Pasha

~

Joshua has a point. However, the post training assessment will only measure knowledge acquisition and NOT utilization to change behavior. The pre- and Post training test will help one gauge the immediate objectives e.g knowledge acquisition by participants on a particular topic, then afterwards as Joshua recommends undertake further monitoring and measurement for intermediate results which will be more on behavior change or utilization of the Knowledge acquired to make a difference.

Alex

~

Adding "Practice" to attitude, skills and knowledge will complete the purpose of the training (ASKP). Actually the grey area that you mentioned/ or internalization of the training is the actual purpose of the training. By pre and post test we can only measure the transfer of knowledge or to some extent the skills. Training need to change attitudes by putting the knowledge/ skills into practice.

Javed Shaukat

~

The discussion is indeed very enriching, but I would still like to pose the same questions that, what would be ideal time, to conduct the evaluation, in order to see the outcomes. Is that something achievable within the project life time?

Sana

~

I recommend you google the work of Donald Kirkpatrick and read about the four levels of training evaluation. He has written a few books and many have written about his work.

Sue

~

With much interest i have been following this discussion on evaluating impact/effectiveness/difference achieved through training/capacity building. we at UNICEF Egypt also are trying to get more rigorous about this.

My understanding is that we should work at more than one level to measure this difference achieved, and at different times. of course we will check on the inputs we used for the training immediately after the training (trainers, funds, material, etc, usually done through workshop evaluations, but also through lecture evaluations if training is a series of lectures). and many have already mentioned the output level, which is basically checking if we have made changes in knowledge and attitude (if appropriate), through pre and post tests.

i would think that we also need to look beyond that and evaluate the difference achieved at outcome and impact level. while impact level is indeed quite far away, and perhaps too much for most of the training, i believe that the outcome level is a must. we want to know whether the training made any difference in the way the trainees work. we could interview them after 6-12 months (and perhaps to a retention knowledge test) as well as their supervisors to see if they are performing any better. we might also interview their customer/clients (for example in the case of health workers, teachers, social workers etc.). this would show whether their change in working has had any impact.

in egypt we have started to use the kirkpatrick 4 level model for evaluations of training (while googling it even appeared on my google bar half way typing "kirkpatrick's four levels of

evaluation"). although much used for business training and thus impact on sales in the end, i think this is very useful to development work as well. and indeed, tools development will be one of the main tasks and challenges. if anyone has anything on these tools, we welcome them very much.
dennis arends

~

In my experience of using the kirkpatrick 4 level (and using social change as an outcome) as an evaluation tool of measuring capacity development effectiveness is that there will be non-training needs (logistics, policies, HR practice) that will affect impact at the community. The method for evaluation should also consider what hinders full effectiveness of training (facilitating and hindering factors)...

Ann lily

~

I agree with Dennis on the use of Kirkpatrick Model as your best reference. Assessing Training effectiveness using the Kirkpatrick as a model has four levels, namely Recall, Learning, Behaviour and Results. It has different evaluation tools complimentary with each other. The tools can be revised to fit your training objectives.

Recall component may be done pre-post type but measures of Learning, Behaviour and outcome would be a challenge for it will lead you to artificial effect that would lead you to embarrassing results. You could be under the Statistical Type I or Type 2 error. If you focus on the Recall component only in your training, then pre-post could be a bet.

Behaviour, attitudes or practices may or may not be improved after training at least within 3 months or the whole lifetime. Just like if you undergo a Legal Training after you studied Law in Harvard University, Recall can be measured outright. But behaviour and practices could not be done outright.

With good post-training support, there were several studies that changes in learning, behaviour and practices as effect of the training initiatives could be observed in 3 months which are sustainable.

Rebecca V Polestico

~

In my limited knowledge, to answer your question is that it depends on how you define outcome. What I learned, there are three levels of outcome: (1) low-level outcome (I prefer to call it, immediate outcome), which by some is considered as no different from output; (2) mid-level outcome, which is **the** outcome; and (3) high-level outcome, which usually considered as "surrogate" of impact. In respect to training, you could 'attach' time dimension to distinguish these division of outcomes. For example, you could evaluate a low-level/immediate outcome (of a training) within 6 month to one year after the training, or evaluate mid-level outcome after 2-3 years (and also the impact, depending the evaluation questions you devise). I assume that the project lifetime is between 3 and 5 years. Last but not least, when you are hired to do an outcomes evaluation, it is important to stick to the definition of outcome provided by your employer to avoid unproductive debate... :-D

In my limited experience, using similar approach I could see the outcomes and impacts of a project that provide post-crisis assistance to recover survivors' livelihood activities in post-Tsunami in Aceh (in which the project lifetime was only one and a half year).

Hope my story helps answering your question.

Thanks & regards,

Ari

~

Thank you for providing examples, I absolutely agree with you.

How about this: tenant-farmers as CBO leaders in power negotiation (change attitude of being submissive for generations) ex-rebel leaders on participatory management; or values re-engineering on government officials.

It is really a dilemma. I remembered arguing the outcome indicators and time frame (for CB M & E substem) that it took us several days to finalized the 1st draft. So we all agreed to emerging impact

– no matter, how small the degree of increase in a limited time-frame (5 years). The need to celebrate with small miracles. As a former colleague said, the former rebel leader not being late in a meeting is also an impact.

Another thing that I find complex is that there are other factors to be considered for capacity building to be effective - systems and procedures, policies, logistics, equipment, education, work experience, etc.

Also to be considered is the similar training provided by other service providers as well. This situation is similar to the micro-credit overlapping, where since it overlaps the % increase in socio-economic well being cannot be attributed to an X project alone.

good luck to you!

Ann Lily Marie O. Uvero

Ann, I agree with you. A capacity building plan or participation planning will indeed avoid us from making mistakes

to conduct unnecessary training or other activities in development program that is not needed by beneficiaries.

What I said about participants may sometimes do not realize yet that the training is actually give them benefit, is in the situation

when a training have a process when the participants' life values and their paradigm need to be changed first before they achieve the skill given by the training

(I think this is the difference between a training given by a profit company to its staff with the trainings for community development).

I actually had experience about this.

I once conduct trainings on Joyful learning methods in class, which also advocate that physical discipline should not be given to students.

The training participants are teachers and school principals that believe that one effective way to teach discipline to students

is by giving physical discipline actions which creates fear and lead to violence towards children.

These teachers and principals born and grew up in a community that believe physical discipline is common things and it is OK to do it,

and this point of view had been lasting for a long time even from their ancestors.

Now, what I'm trying to say is that they do not realize yet that non physical disciplines can actually be a more effective strategy

in teaching students in class, since it creates a more fun atmosphere for students learning in class --- because they don't have evidence yet.

They need to practice this new knowledge first, find the proof, then the behavior change, and I say it will take a long time, even if it is supported by post training activities for the participants such as on the job trainings, small workshops, Re-fresher trainings, and so on until the desired outcomes achieved.

Perhaps my conclusion is, a training for increasing participants knowledge & skill without a process of changing one's life values and paradigm,

the outcomes will likely easily to be achieved, rather than the training which require life values and paradigm changes.. the outcomes will likely harder and longer time to be achieved, and from this we can determine when is the suitable period for us to conduct an evaluation towards training effectiveness.

Laura

Dear Sana,

<it is important to stick to the definition of outcome provided by your employer to avoid unproductive

> debate... >

Yes, I have learned the hard way that the employer or contractor calls the shots, even if he's supposed to be the evaluand. However, arbitrary time dimensions don't change the nature of phenomena. You can call something an impact, or a "surrogate" (proxy) for an impact, but that doesn't mean it is really helping people or the environment, just because you measure it a few years later.

In Brazil in the nineties, putting "institutional development" funding from World Bank or IADB loans into a UNDP (later, also Unesco) project to do "capacity development" often became an excuse for doing nothing but hiring "consultants" to do the work of bureaucrats, importing computers through UN channels, travelling and providing "training." Sometimes program evaluations done years later could find no evidence whatsoever that Health Ministry projects had improved health (or

even attempted to do so), Education Ministry projects had promoted schooling (much less student learning), or Environment Ministry projects had improved (or even protected) the environment. Actually, evaluators were sometimes strongly discouraged from even looking for such impacts. Suggestion: if an initiative is intended to do nothing but help the learners themselves learn, let's not call it training; let's call it education.

Robert

Dear Vardhani,

Here is my suggestion.

I don't know how frequent the training is being provided to the teachers. But, let assume there are 3 phases of training in one year, and they will be done every 4 months. I'd prefer training's phases/stages rather than a series of training if they are meant to build the capacity of beneficiaries gradually. Say, every phase is a prerequisite for the next and every participant must reach a pre-defined level of practice/capacity (i.e. outcomes) in order to continue to the next phase. In this case, you could assess/evaluate the effectiveness 2-3 months after one phase, or a couple of weeks before the next phase is carried out. Of course you need to carefully define measurable indicators for outcome of each phase – in your case probably are knowledge, attitude and practice/skills. The challenge with this approach is that when the number of participants who goes to the next phase is gradually decreasing in contrast to increasingly difficult to grasp materials. You need to identify immediately (of course through formative M&E) what causes this – whether the design of the training, too high expectation, weak human resources, the way training materials are being provided, etc. – in order to keep the project afloat. While this approach is said to be highly adaptive (theoretically), in practice it is no secret that not easy to change project work plan & direction once it starts. In my opinion, there should be a continuous post-phase accompaniment (mentoring) in order to help those who did not pass a training phase, so in the end all the participants will eventually continue to the next phases and graduated with expected knowledge & skills. But, I am not sure whether this could be done easily as most project-based training often taking assumption that it is designed perfectly and participants do not need to repeat it... :D

Hope my story helps you.

Regards,

Ari...

Hi,

In measuring training effectiveness, another critical dimension to discuss is whether it is a one-off training or a series of trainings for the same group. For eg: in our projects we conduct trainings for school teachers. Each teacher undergoes a series of trainings to ensure that they have adequate capacity to handle a disabled child. These trainings are spread throughout the project lifetime. In such a case, how does one measure the effectiveness and at what frequency?

Vardhani

Robert and others,

I would think that it actually does not matter that Kirkpatrick was developed for the business sector and training for business improvement (sales?). I feel that in the non profit sector we should also be concerned with business improvement; that is, if we train people (build their capacity), I assume we do this with the aim that the trainees will do their job better. and eventually we improve the lives of those we work for (poor, children, women, vulnerable groups, etc.).

I have not thought about whether this model will actually be helpful for elementary education (I doubt, and I think that educators have a host of tools measuring impact of teaching in schools), and would think it is mainly for adult/working people training, of which we do quite a bit at UNICEF, hence the interest in measuring the impact.

Several people have also raised the issue of attribution; how can we control for factors beyond the inputs of the training itself (incl. the quality of the facilitator)? but here I would think that at the output level we should certainly be able to claim result (or failure). We are the ones to recruit and select the facilitators, we are the ones that can ensure certain curricula and training circumstances. Of course at a higher level than that, attribution plays a role. but the point is that many of us do capacity building as an activity and an expected result/end (very much the human rights based approach to development cooperation)--so, we should consider evaluation of this intervention as any other project/major intervention. and as such attribution comes in, no escape indeed. but is this not same for any development project?

Cheers, Dennis

9. M&E News Summary
February 2009

Mapping exercise of all Civil Society Organization

Dear all,

I need help in writing a brief concept of about 11/2 pages on how I view a mapping exercise of all CSOs operating in the field of HIV and AIDS in Swaziland which I am applying for as a regional enumerator. This mapping exercise is part of an Internal Organizational Assessment commissioned by UNAIDS through the Coordinating Assembly of Non-Governmental Organizations in Swaziland (CANGO).

Please help me out on the structure of this brief and what I write about. The deadline is **4th February 2009**. Help me out colleagues ASAP.

Regards
Sandile Ginindza

Hi Sandile,

How many organizations are you wanting to map. Are you wanting to map them geographically or against the outcome areas they are attempting to influence?

Regards,

Paul Duignan, PhD
paul@parkerduignan.com

Dear Sandile,

You might want to look at Net-Map (<http://netmap.wordpress>) as a tool for mapping your civil society organisations. It's a low tech method for mapping actors, formal and informal links, goals or actors and their influence on the issue at stake. Have a look at the website for a manual and case-studies that might help you develop your own strategy.

Kind regards
Eva Schiffer

Hi again Sandile,

I realize that you have a tight time frame, so I have not waited for a response from you on the number of projects involved. I have done what I am describing below with up to 30 projects, but it could be done with more.

Below is an example of how you can map multiple projects onto an outcomes model if that is what you want to do. You build the outcomes model according to the 13 Tips here <http://easyoutcomes.org/files/13tipsa4-205.pdf>

And then you can map your projects onto the common outcomes model. (you could map them onto various common models you can develop, including HIV related outcomes, outcomes related to the nationwide system dealing with HIV, specific population groups etc.)

A mock-up of what this looks like is available

at <http://www.outcomesmodels.org/models/educationsector39.html>

Just click through all of the pages (use the green down arrow at the top of the page) and you will find an example of how the projects can be mapped onto the model.

And you can see here how you can put in the number of projects that map onto each outcomes to assist strategic discussions in the example below <http://www.outcomesmodels.org/models/communityprojectmapping29.html> These examples are web page models which have been created from within DoView outcomes and evaluation software. The first example is mocked up in a soon to be released no cost update version which will allow images to be included. The second example is in the current version of DoView. You can download a trial version of DoView from <http://www.doview.com/download.html> and you can download the second example by clicking on Download the DoView file of this model on the bottom of the screen when you are looking at the second example. You can get other models to play around with at <http://www.outcomesmodels.org>. You can get a full workbook on how to do all this from http://www.easyoutcomes.org/files/easyoutcomesworkbookv1-9sv_2008-4-11.pdf If you need any help, just let me know.

[Disclosure: I have been involved in the development of DoView software]

Regards,
Paul Duignan, PhD

Hi Sandile and all

I forgot to mention, that there is an example of a NGO network mapping exercise available online, here:

[Davies 2007](#) "Research and Advocacy Organisations and their relations with issue-base coalitions in Ghana" report on an online survey

at <http://www.mande.co.uk/ghana/Online%20survey%20of%20coalitions.doc>

regards, rick davies

Citizens Report Card and Exit Polls

Dear All,

I am a new member of the group and have within this very short time benefited a lot from senior colleagues in the field of M and E. My current task is setting up an M and E systems for a governance programme in Sierra Leone. Our focus is to link organisation working to create the space, and build meaningful communication, between citizens, civil society organization and state. The programme log frame and M and E Plan has been finalised and agreed by our donors. There are a number of tools I hope to develop for effective and routine monitoring of our interventions.

I am writing to seek information from members of the group to share the following tools and training manual with me:

1. Citizens Report Card – For evaluating of citizens satisfaction with basic service at the community level. We are looking at Access, Utilization and Satisfaction with service at the community level. Also we want to evaluate the relationships and interaction between local councils and their communities.

2. Exit Polls – For evaluating dialogue sessions, training and solidarity events

Thank you very much for your support.

Patrick Elogima Robin

Dear All,

Can any one in the group help me out with citizen report card and exit polls evaluation techniques. I am currently working on a governance programme and will want to evaluate citizens satisfaction with the quality of services and thier interaction with local councils.

We also want to evaluate solidarity events and training programmes

Patrick Elogima Robin

Dear Robin and Ann,

You may also want to take a look at the Affiliated Network on Social Accountability (ANSA) Africa website (<http://www.ansa-africa.net>) for resources on social accountability and to link to a large network of social accountability initiatives and actors.

The secretariat of this network has recently moved from the Human Sciences Research Council (<http://www.hsrc.ac.za>), a large public entity in South Africa to IDASA (<http://www.idasa.org.za>), which is a strong South African NGO involved in governance related matters.

Both organizations have experience with Social Accountability methods.

Kind regards,

Marlene Roefs

Interim project manager

ANSA Africa – Democracy and Governance Programme

Human Sciences Research Council

South Africa

Dear All,

An Email based discussion group called Solution Exchange is supported by the United Nations in India. A Consolidated Reply on this discussion group is a good resource for developing Human Development/ governance Report Cards by citizens. I am giving below the web link to this document. Please contact me if you are not able to access this document for any reason:

<http://www.solutionexchange-un.net.in/decn/cr/cr-se-decn-07120701.pdf>

Hope this is helpful to Patrick.

Thanks

Pankaj

Dear Robin,

Am in the field right now conducting a mid-term review for one of our HIV/Aids projects. I am currently using a modem and the download/upload speeds are pretty slow. I will send you the design template and report of a CRC survey that I was one of the lead persons for on Govt service delivery in Kenya on Monday when I get back to the office on Monday. I hope that will not be too late.

Thanks

Lewis Aritho

Liverpool VCT

Nairobi, Kenya

Question on Framing Evaluation Questions

Hi,

I just wanted to get your insight on the best way for framing evaluation questions. Is it appropriate to base the summative evaluation (and take evaluation questions) directly upon the project/program objective statements and the scope of work as defined in the contract/proposal/any official documentation?

If so, then theoretically what type of evaluation is that? (e.g. evaluation serving accountability only? or ...)

Look forward to your comments,

Sana

M&E research coordinator, Pakistan

Dear Sana

Let me allow to ask you about purpose of evaluation before my cement. Do you want evaluation for project process, output evaluation or impact evaluation? .Usually Project

proposal should have its M & E plan and on the basis of that one can plan evaluation. Evaluation is usually done for process/output /result or may be impact. It depends upon researcher if he/she wants to evaluation process, output, result or impact. Some time objectives statement are too widened or ambitious and its difficult to frame out on objectives which may take more time to achive in contrary to project timeframe.

Learner

Javed Pasha

Senior Livelihood Officer

FAO UN Pakistan

Hi All,

The intended user of the evaluation i am referring to is the management of the center, running the project and the donor agency. The intended use of the evaluation findings is to learn about the project outcomes as a test-case for scaling up on similar initiatives (designing similar initiatives as pointed out by Mr. Williams) and from the donor point of view it is to serve accountability purpose.

The center's management is interested on a kind of output evaluation and to some degree impact aswell, but really the focus is to see what has been achieved in terms of the objectives that were set.

Therefore in such a case I would request suggestions that is it alright to start by taking the project objectives (both general and specific objectives and the planned scope of work, but define more specific sub-questions) for evaluating if the project has effectively achieved the set goals or not.

Best,

Sana

Some additional M&E Resources

M&E Short Cuts and Field Friendly Modules

CRS (Catholic Relief Services) and American Red Cross recently completed a series entitled M&E Short Cuts. Each edition is a few pages in length, and is available in English, French and Spanish.

The topics covered are designed to respond to field identified needs for specific guidance and tools that did not appear to be readily available in existing publications. Program managers as well as M&E specialists are the intended audience.

The Short Cuts series provides a ready reference tool for people who may have used the full modules (French and Spanish versions to be posted shortly), those who simply need a refresher in the subject, or those who want to fast-track particular skills.

The topics covered include:

- 1 Preparing for an Evaluation
- 2 Managing and Implementing an Evaluation
- 3 Reporting and Communicating on an Evaluation
- 4 Writing Human Interest Stories for M&E
- 5 Monitoring and Evaluation Planning
- 6 Using Indicator Performance Tracking Tables
- 7 Hiring M&E Staff
- 8 M&E and Ethics
- 9 Capacity-Building Guidance
- 10 M&E Planning

Both the modules and the Short Cuts series can be found at:

<http://crs.org/publications/list.cfm?sector=19>

and

<http://www.redcross.org>

Both the Short Cuts series and the longer modules were produced under

the respective USAID/Food for Peace Institutional Capacity Building Grants to CRS and American Red Cross. The views expressed are those of the authors and do not necessarily represent those of the USAID or FFP. Although both the Short Cuts series and longer modules focus on Title II programming, it is hoped that the both will have value beyond just the food-security realm.

Reflective Peacebuilding: Planning, Monitoring and Learning Toolkit
CRS and the Joan B. Kroc Institute for International Peace Studies has also produced a toolkit on M&E for those engaged in peacebuilding initiatives.

This is available at:

<http://crs.org/publications/list.cfm?sector=8>.

Undertaking M&E-related activity in a consortium

As a collaborative initiative, The Consortium Alignment Framework for Excellence (CAFE) assists agencies to improve their ability to form strong consortium that respond to the needs of the most vulnerable and meet donor requirements while strengthening the relationship between them. The focus of CAFE is to make the managerial, financial, and administrative functions of a consortium effective, efficient, and supportive of project goals, community needs, and donor intent. CRS staff and partners working in consortium, with input from headquarters staff with consortium responsibilities developed CAFE.

This document is available at:

<http://crs.org/publications/list.cfm?sector=23>

Kind regards,
Guy Sharrock

Measuring Capacity

Dear Mushtaq,

Yes, you are right. Capacity must be looked at individual level too. UNDP's capacity assessment tool attached to my previous email clearly states that the capacity resides in three regimes:

1. Individual level
2. Organizational level
3. Enabling environment

Well, KAP assessment of the trained people could be good but the application of knowledge again depends on three main factors as mentioned above. I hope the following website will be helpful for more information on performance assessment at individual levels.

<http://careercompass.berkeley.edu/perfmgmt/resources/evalforms.html>

Thak you
Bhandari

School and Hospital Evaluation

Dear Colleagues,

I am Saad Yousaf and working as a Team Leader for Islamic Aid (a UK based charity). We are about to conduct an evaluation of our running chain of schools and hospitals in Pakistan. Furthermore, to conduct a results-based performance evaluation framework of hospital and teaching staff. Just wondering if I can have access to the already available resource (outlines or guidelines) on these topics or somebody can share the reports they already

have which can guide us in doing the task with professional approach and methodology. The size of schools and the hospitals is relative small but providing quality services. Many thanks.
Saad Yousaf.

Meta-analysis

Dear KG,

There are three concepts related to your question. These include meta-evaluation, meta-analysis and evaluation synthesis.

Meta-evaluation (sometimes called pooling evaluation) is the evaluation of the quality of evaluations using various standard criteria. Some times it is also known as 'the evaluation of evaluations'. D.Stufflebeam (1974) defined meta-evaluation as a procedure for describing an evaluation activity and judging it against a set of ideals concerning what constitutes good evaluation. In other words, meta-evaluation means assessing the merit of evaluation efforts. Similarly, Michael Q. Patton (1997: 143) defined meta-evaluation as a process to evaluate a given evaluation studies based on the profession's standards and principles. The motive for conducting meta-evaluation in Patton's terms is to ensure an independent and credible review of an evaluation's strengths and weaknesses. In this meta-evaluation context we are then dealing with questions such as: was the evaluation well done? Is it worth using? Did the evaluation meet professional standards and principles? and so forth (see also Patton, 1997: 193, 333).

Meta-analysis is a statistical way of aggregation of the results of individual studies. Glass et al (Glass, 1976) identified three characteristics of meta-analysis. First, meta-analysis is quantitative and uses numbers and statistical techniques for organizing and extracting valuable information that is nearly incomprehensible by other methods. Second, meta-analysis does not tend to evaluate the quality of existing studies. However, meta-analysis attempts to record various aspects of research methodologies for the existing studies to identify their relationship to study findings. Third, meta-analysis aims to compare existing studies and to seek general conclusions across studies.

An **evaluation synthesis** is a "content synthesis of various evaluation studies (largely qualitative)", i.e. a global evaluation/cross-section analysis (Widmer, 2003). An evaluation synthesis requires a thematic clustering among the evaluation studies upon which it is based, but has a less narrow focus due to its qualitative approach.

It is important to note that mostly meta-evaluation precedes meta-analysis. This is because evaluations should be included in to a meta-analysis only when they fulfill certain level of minimum quality standards. Hence, based on your explanation you are in need of conducting meta-evaluation (to evaluate which evaluations are good/poor) and meta-analysis (if the evaluations are quantitative) or evaluation synthesis (if the evaluations are largely qualitative).

I hope this helps some.

With best regards,

Hailemichael Taye
Monitoring and Evaluation Specialist
Agriculture Sector Support Project(ASSP)
Addis Ababa, Ethiopia.

Request for Independent Evaluation Form

Dear Colleagues:

The Government of Liberia and her development partners have initiated a project called the "Senior Executive Service" (SES) which purpose is to recruit Liberian professionals to work for the government in various ministries and agencies on performance contract basis. One of my immediate function as M&E Officer is a to develop a form that will be used to evaluate SES staff that are deploy in the various agencies. This form will be used by

independent evaluators i.e. individuals that are working in those agencies but does not have direct oversight responsibility of SES staff.

As per the contract document, individuals should be awarded if they are performing while those not performing will also not be rewarded. To avoid individual supervisor from writing misleading/negative information (so to speak) on the SES staff for whatever reason, we think it is necessary to develop an independent evaluation form as mentioned above.

I am therefore asking if anyone in this group can assist me with such a form or can give me a link where I can possibly find a document that can guide me in the process. If there is anything that is being currently use in any of your agency that I can use as a guide will be highly appreciated.

Thanks in advance.

Isaac

Dear Isaac,

Tool development is one of the essential parts of any evaluation. If you have to evaluate the efficiency of staffs through an unbiased tool, first of all you have to give operational definition to the term •efficiency•. You also have to decide whether the efficiency requirements are same for all level staffs. You can develop the definition through your project objective as well as from the roles& responsibilities description of the staffs.

While preparing the questions/ checklist (items) you have to develop 50% more items. For example if you plan to develop a tool having 100 questions, you have to prepare an initial list of 150 questions. After pilot study, while analysing the results you have to give many considerations for short listing the items. Items have no discriminative power (getting same scores for all individuals), high difficulty level etc should be eliminated. There are some statistical analyses to find out the reliability of the tool. If you read more on •Item Analysis• (mostly for development of psychological tools) you get more insight.

All the best for your endeavour!

Thanks & regards

Sunitha Thampi, PhD

MIS & Data Management Officer

Project Concern International

Pune

How to estimate sample size for control group

Dear All,

I joined this group two days ago and am pleased to contribute and share my skills.

There are various methods of calculating your sample size. However there are issues that you need to have in mind, these includes Level of precision commonly referred to as sampling error, the level of confidence and the degree of variability or what you would call distribution of your study attributes in the population. Remember the more heterogenous is your population then a larger size would be necessary to obtain the given precision.

Purity Njagi

Monitoring and evaluation Officer

On Tue, Feb 3, 2009 at 7:03 AM, jalandhar pradhan

<jpp_pradhan@yahoo.co.uk> wrote:

> Dear All,

> We are planning to conduct a baseline survey for an intervention project.

> The study is based on case-control design.

> So, could you please suggest me, is there any methodology to estimate the

- > number of sample/ village to be covered for control area for comparison
- > purpose.
- > Infact for project area we are estimating the sample size using the 'P1' and
- > 'P2' value.
- > Can we apply the same formula to estimate the sample size for control area?
- > Waiting for your valuable suggestions.
- > Regards,
- > JAlandhar

There are lots of methods for sample size calculation, but it depends on how sophisticated you want this to be. Least sophisticated is to assume fixed treatment effects throughout, in which case you can just use the usual sample size calculations to determine the overall N needed, and then allocate shares of the sample over the clusters (in this case, villages) in a manner that either optimizes precision or streamlines analysis by minimizing the use of weights (there's a tradeoff). In this case, overall sample size can be computed using standard sample size calculation software, such as the `sampsi` function in Stata. Precision or "streamlining" as above requires application of some basic sample design principles.

But if you think there may be some heterogeneity in treatment effects or that there may be random effects associated with clusters, then you would want to use some form of cluster design framework to determine optimal number of clusters and optimal number of units to be sampled within each cluster. There is a program called "Optimal Design" that performs these calculations:

<http://sitemaker.umich.edu/group-based/home>

Or, you can set up a custom simulation-based program to determine sample size. I've done this for some of my evaluation work. If you are interested, we can talk details.

Cyrus Samii

Political Science

Columbia University

cds81@columbia.edu

Burundi Survey: www.columbia.edu/~cds81/burundisurvey/

ISERP Statistical Consulting:

www.iserp.columbia.edu/services/statistical_consulting.html

Comparative Political Economy Blog: cpecolumbia.blogspot.com

Greetings from snowy Connecticut--

Sample size calculations are the same, regardless of whether it is for the treatment or control group. Sample size is the function of the probability of the occurrence of the event, and the desired precision of the estimate. If you need to make comparisons between 2 groups, you are going to want relatively precise estimates otherwise it will be difficult to show that the treatment had a statistically significant impact on the population of interest.

Beth

Qualitative Evaluation of National Development Plans

Hello all

i was just wondering! has anyone ever done a qualitative evaluation of national development plans? i mean documenting benefits accrued to catchment populations? further making an analysis of what people think of quantity indicators as they relate to the improvement of their lives? for example, what do built schools translate into as far as

beneficiary communities are concerned? does it matter if the NDP has a GDP target of some amount to the common man?

i will be glad to hear from you

Mapoma Christopher
University of Zambia

There is a long list of universities that offer programs related to evaluation. You can see them at http://www.eval.org/Training/university_programs.asp.

Below is a website of the American Evaluation Assoc. with a list of universities offering degrees in M&E. I am currently here in Florida as a Fulbright visiting prof and I can tell you it's indeed enriching and worth the effort. And you got the right choice of degree to pursue -carry on...

http://www.eval.org/Training/university_programs.asp.

All the best!

Romeo Santos

Dear Mapoma,

you are raising a very weighty issue. As far as I am aware, there are two main types of evaluations/impact assessments related to national development plans.

One is an ex-post vs. ex-ante comparison of a range of indicators linked to the planned objectives. Such types of analyses have often been done, and tend to mainly be of a quantitative nature, and based on types of information systems that exist. In India, for example, we have a very wide data base and ongoing monitoring that provides the information for such types of evaluations. In Zambia, some such measures may also be available, and could be used to assess the extent to which the 'common man' or woman may have benefited.

Another approach could be to take specific policies within the development plan and evaluate them. This becomes a much more complex analytical exercise (in my opinion), as the externalities and interactions are much harder to unravel.

However, your question is more in lines with a qualitative analysis, that is almost of a political economy type of investigation.

Would be intersted to hear what others in our community have to say on this.

Very best wishes,

Shubh

(Ms. Shubh Kumar-Range)

Evaluations for Human Development

New Delhi

Pre and Post Test for Measuring Training Effectiveness

Dear Colleagues,

I hope you can help me on this issue.

I am planning to conduct trainings, and thinking to use pre and post test as one of the tool to measure the effectiveness of the trainings.

However recently I read in a Monitoring and Evaluation manual, it is not recommended to use pre and post test as the tool to measure the training effectiveness, I'm not quite understand why, since there is no argument to support this statement.

Maybe there are any of you that have experience in using pre and post test to measure the trainings and can explain to me why it is not recommended. I also would like to know if there is another tools I can use to measure the training effectiveness.

Appreciate your sharing.

Cheers.

Laura Ginting

Development worker

Hi Laura,

One of the reasons behind why Pre-Post Test design are not advisable to be used for training effectiveness because of "intervention time". Changes that may take place because of training delivery would be extremely artificial at that time of measure. And it will lead you to conclude that the training is effective and we knew it is NOT. For Pre-Post design, intervention time is an element of primary importance.

However, at some stage, the Feedback of the Training would be more appropriate which also depends on your evaluation instruments.

Sincerely,
Becky P

Question - when you say you want to measure the effectiveness of training, what exactly do you want to measure? Whether, for example, people learned something from the training, or whether they took the training and did something with it.

In either case, pre and post test seems the best way to determine any effects of training, along with all the other issues of random assignment, self selection, etc. that you also have to deal with.

Gene

Dear Gene,

Relating to your question, As far as I know, Pre and Post Test is used to measure the training outputs, that is to know whether the training participants had gain the knowledge given in the trainings.

So, I guess we usually applied the tools in the first day of training (for pre test) and at the last day of the training (for Post test) - *am I right? Please do correct me if I'm wrong.*

When I read the statement that says "The test is not recommended to be used for measuring the trainings" I assumed the person stated that is not recommending the M&E officers

to use it for measuring the training outputs, and it makes me to ask Why?

But from your opinion and other colleagues opinion, I can make conclusion that Pre and Post Test is still the best tool for measuring the training outputs (is the trainee has gained the knowledge given from trainings),

and it still is the best as long as we do not use it for evaluation (for outcome, and impact). -

Please also correct me if I make the wrong conclusion.

Thank you for all the replies, your opinions had enriched me.

Laura Ginting

Development worker

Yes, "Pre and Post Test is still the best tool for measuring the training outputs (has the trainee has gained the knowledge given from trainings)"

At least at one particular time. As Becky also wrote "Changes that may take place because of training delivery would be extremely artificial at that time of measure."

That is, right after the training ended, people who do better appear to have learned the material, or have learned how to respond to way you want them to respond.

However, whether this learning translates to doing anything with it, after the training, that is another question, and you won't get an answer to that question by giving a survey right before training and right after training. You would need another survey or some kind of feedback later on, after people return to their 'home' settings, and had a chance to apply what they 'learned'.

Hope this helps.
Gene Shackman

Laura,
Pre post approach can judge the delivery of training and NOT its effectiveness. The effectiveness of the training can be measured after the training is applied by the training recipients/ or the effect that the training bring to the organization (depending on the type of training).

Shaukat
Hi Laura,
I want to support what Shaukat has said. You can use pre and post test to measure training delivery and what new information the participants have acquired in the course of the training.
However the effectiveness of the training itself will take time. This will best be measure after sometimes, you can then measure if the training as produced the desired improvement in the delivery of the participants work with an eye on the objective of the training.
World Bank has a whole document on measuring the effectiveness of their training. the document is titled " Using training to build the capacity for development" you can download one or request for a copy.
Have a pleasant training
Olanipekun Oluwasola G. (Nehi)
Monitoring and Evaluation Manager
Life Vanguard, P.O.Box 2182
Osogbo, Osun-State. Nigeria

Dear Laura,
It is really an interesting discussion. However, all said and done, I would still go with pre and post evaluation. But we must be very clear what indicators we want to measure. For example if you want to measure the gender attitude among the trainees, you may have a pre training evaluation and post training evaluation and if you want to see the sustainability of the training, you may still want to to another follow up evaluation after certain period.

In any case, pre and post test evaluation is the most effective methods to measure effectiveness, either in change of behavior or in attitude, but we should be very clear what we really want to measure.

I hope that is helpful.

Ajay K Singh
International Centre for Research on Women (ICRW)
C-139 Defence Colony
New Delhi- 110024

Hi Shaukat,
But then my question would be that is it something measurable or evaluated "near" to the project end or soon after it. How do we see such kind of effectiveness if the project evaluation has to be done within one month of the project end (that includes the report writing time also).
Best,
Sana

Dear Shaukat,
Much as it may be too early to ascertain effectiveness of a training just after a few months, as M&E practitioners, through our monitoring I think we need to focus a little more on results of activities such as trainings. This is what Results Based Management (RBM) is all

about. Instead of just looking at outputs of trainings, we can seek to establish through monitoring whether actually these trainings are having the desired results. By doing this, I think we can later on (during evaluation) be able to attribute certain changes in the project environment to our activities (such as trainings).

Thanks,
Titus G Tumusiime
M&E Officer,
GOAL Uganda

Dear all,

There is a better way of assessing training sessions and measuring knowledge acquired (output) by trainees in the course of participation in training. Though, it's still pre-post test but this is administered on each day of training over presentations for that day.

Methodology

M&E officer will ask for at least five questions from each facilitator each day which will be administered on the participants before the training and after the training daily. On each day, previous daily performance should be displayed in a chart to show increase/decrease in knowledge on each sessions presented. The approach is actually tasking but will help to measure the following:

- Facilitators' skill in disseminating topic s/he planned to.
- Participants' knowledge gained/lost.
- Its makes participants actively participate in training sessions knowing that at the end of each day there will be test.

We should also note that if there is difficulty in measuring immediate result (output) of inputs, measuring outcome may be impossible. Therefore, we cannot conclude that "Changes that may take place because of training delivery would be extremely artificial at that time of measure," as Becky wrote.

We should also bear in mind that •MONITORING• is a continuous activity, after training people for a particular activity we should device means (tool) for ensuring that learning translate to doing something with it.

Regards

AWOLEYE JOSHUA OLATUNJI
MONITORING & EVALUATION COORDINATOR
SOCIETY FOR FAMILY HEALTH
NO. 8 PORT HARCOURT CRESENT,
OFF GIMBIYA STREET, AREA 11 GARKI
P.M.B. 5116 WUSE
ABUJA

Sana,

The achievement of results is gradual i.e from input-output-outcome-impact. It is possible to see the outcome of your efforts within the lifetime of the projects which starts building up after achievement of the outputs within certain assumptions. That is if you are able to see the immediate results given certain assumptions then there is a likelihood of seeing the intermediate results and the outcome within the lifetime of the project. The duration of the project also matters and the kind of intervention being undertaken but it is possible to see outcomes being generated slowly by slowly as time elapses. Just think a bout process of making wine. The wine matures with time.

Alex

Hi Sana,

When it comes to evaluate training outcome, I guess there is no particular ideal time for it, since in my opinion the more participants feel the benefit from the training they've got

(the given training is exactly what the participants' need),
the behavior will change faster means the expected outcome will happen faster,
and it is possible to be achieved within the project lifetime.
The problem is, sometimes we give training that is not needed by participants or they did not realize yet that
the trainings will actually benefit them maybe in the future. This will cause the behavior changed slowly,
means the expected outcome will also slowly to be achieved, sometimes took years.
In this case, I have experience to make evaluation of the trainings at least in 6 months even though the outcome had not achieved yet,
but the process towards achieving the outcome sometimes can be identified.
Besides, usually donors would required to have evaluation report in every 6 months.
Perhaps others can discuss this too....

Laura

Dear All,

Activities (like trainings) can be measured monthly. If we talk about training outputs (ie. knowledge gained from training), I don't think we can measure only at post-test, as knowledge tends to decline over time. When I had worked in Vietnam, we did post-tests, but then re-tested a sample of training participants every 6 months to see if there was still good understanding. Often it was not, which forced us to re-design activities (often by doing refresher trainings, improved trainer-quality, reducing #s of training participants, etc.). Since our community development work was done in 15 years chunks, with impact evaluations done every 5 years, we measure outcomes (ie. changes in behavior) every 5 years. However, I came to the conclusion that this was not sufficient and that some outcomes could be measured annually. We did not want to wait 5 years to see if knowledge actually translated into behavior.

Greg

I think Javed has rightly added word Practice to ASK complementing the purpose of training but how effective could be training in term of Practice, Knowledge, Attitude and skill are totally different things to be measured in context of training effectiveness. How effective the training was delivered is something else depends upon its contents, methodology and audience understanding or its acceptance. Training effectiveness in term of practice or attitude can't be judged in pre and post training evaluation . It can't be easy to measure effectiveness of training in attitudinal change or practice of people right after completion of training as it would take time to have attitudinal change

kind regards

Mr. Javed Iqbal Pasha
Senior Livelihoods Officer
FAO UN

Joshua has a point. However, the post training assessment will only measure knowledge acquisition and NOT utilization to change behavior. The pre- and Post training test will help one gauge the immediate objectives e.g knowledge acquisition by participants on a particular topic, then afterwards as Joshua recommends undertake further monitoring and measurement for intermediate results which will be more on behavior change or utilization of the Knowledge acquired to make a difference.

Alex

Adding "Practice" to attitude, skills and knowledge will complete the purpose of the training (ASKP). Actually the grey area that you mentioned/ or internalization of the training is the actual purpose of the training. By pre and post test we can only measure the

transfer of knowledge or to some extent the skills. Training need to change attitudes by putting the knowledge/ skills into practice.

Javed Shaukat

Dear All,

The discussion is indeed very enriching, but I would still like to pose the same questions that, what would be ideal time, to conduct the evaluation, in order to see the outcomes. Is that something achievable within the project life time?

Sana

I recommend you google the work of Donald Kirkpatrick and read about the four levels of training evaluation. He has written a few books and many have written about his work.

Sue

dear all,

with much interest i have been following this discussion on evaluating impact/effectiveness/difference achieved through training/capacity building. we at unicef egypt also are trying to get more rigorous about this.

my understanding is that we should work at more than one level to measure this difference achieved, and at different times. of course we will check on the inputs we used for the training immediately after the training (trainers, funds, material, etc, usually done through workshop evaluations, but also through lecture evaluations if training is a series of lectures). and many have already mentioned the output level, which is basically checking if we have made changes in knowledge and attitude (if appropriate), through pre and post tests.

i would think that we also need to look beyond that and evaluate the difference achieved at outcome and impact level. while impact level is indeed quite far away, and perhaps too much for most of the training, i believe that the outcome level is a must. we want to know whether the training made any difference in the way the trainees work. we could interview them after 6-12 months (and perhaps to a retention knowledge test) as well as their supervisors to see if they are performing any better. we might also interview their customer/clients (for example in the case of health workers, teachers, social workers etc.). this would show whether their change in working has had any impact.

in egypt we have started to use the kirkpatrick 4 level model for evaluations of training (while googling it even appeared on my google bar half way typing "kirkpatrick's four levels of evaluation"). although much used for business training and thus impact on sales in the end, i think this is very useful to development work as well. and indeed, tools development will be one of the main tasks and challenges. if anyone has anything on these tools, we welcome them very much.

cheers,

dennis arends

chief, social policy, monitoring and evaluation

unicef egypt

Dear all,

In my experience of using the kirkpatrick 4 level (and using social change as an outcome) as an evaluation tool of measuring capacity development effectiveness is that there will be non-training needs (logistics, policies, HR practice) that will affect impact at the community. The method for evaluation should also consider what hinders full effectiveness of training (facilitating and hindering factors)...

With best regards,

Ann lily

Hi Laura,

I agree with Dennis on the use of Kirpatrick Model as your best reference. Assessing Training effectiveness using the Kirpatrick as a model has four levels, namely Recall, Learning, Behaviour and Results. It has different evaluation tools complimentary with each other. The tools can be revised to fit your training objectives.

Recall component may be done pre-post type but measures of Learning, Behaviour and outcome would be a challenge for it will lead you to artificial effect that would lead you to embarrassing results. You could be under the Statistical Type I or Type 2 error. If you focus on the Recall component only in your training, then pre-post could be a bet.

Behaviour, attitudes or practices may or may not be improved after training at least within 3 months or the whole lifetime. Just like if you undergo a Legal Training after you studied Law in Harvard University, Recall can be measured outright. But behaviour and practices could not be done outright.

With good post-training support, there were several studies that changes in learning, behaviour and practices as effect of the training initiatives could be observed in 3 months which are sustainable.

Hope this helps

Rebecca V Polestico

Unit 803 The Asia Tower

151 Paseo de Roxas corner Benavidez Sts

Makati City, Makati Metro Manila

1223 Philippines

Dear Sana,

In my limited knowledge, to answer your question is that it depends on how you define outcome. What I learned, there are three levels of outcome: (1) low-level outcome (I prefer to call it, immediate outcome), which by some is considered as no different from output; (2) mid-level outcome, which is **the** outcome; and (3) high-level outcome, which usually considered as "surrogate" of impact. In respect to training, you could 'attach' time dimension to distinguish these division of outcomes. For example, you could evaluate a low-level/immediate outcome (of a training) within 6 month to one year after the training, or evaluate mid-level outcome after 2-3 years (and also the impact, depending the evaluation questions you devise). I assume that the project lifetime is between 3 and 5 years. Last but not least, when you are hired to do an outcomes evaluation, it is important to stick to the definition of outcome provided by your employer to avoid unproductive debate... :-D

In my limited experience, using similar approach I could see the outcomes and impacts of a project that provide post-crisis assistance to recover survivors' livelihood activities in post-Tsunami in Aceh (in which the project lifetime was only one and a half year).

Hope my story helps answering your question.

Thanks & regards,

Ari...

Laura,

Thank you for providing examples, I absolutely agree with you.

How about this: tenant-farmers as CBO leaders in power negotiation (change attitude of being submissive for generations) ex-rebel leaders on participatory management; or values re-engineering on government officials.

It is really a dilemma. I remembered arguing the outcome indicators and time frame (for CB M & E substem) that it took us several days to finalized the 1st draft. So we all agreed to emerging impact – no matter, how small the degree of increase in a limited time-frame (5 years). The need to celebrate with small miracles. As a former colleague said, the former rebel leader not being late in a meeting is also an impact.

Another thing that I find complex is that there are other factors to be considered for capacity building to be effective - systems and procedures, policies, logistics, equipment, education, work experience, etc.

Also to be considered is the similar training provided by other service providers as well. This situation is similar to the micro-credit overlapping, where since it overlaps the % increase in socio-economic well being cannot be attributed to an X project alone.

good luck to you!

Ann Lily Marie O. Uvero

Ann, I agree with you. A capacity building plan or participation planning will indeed avoid us from making mistakes

to conduct unnecessary training or other activities in development program that is not needed by beneficiaries.

What I said about participants may sometimes do not realize yet that the training is actually give them benefit, is in the situation

when a training have a process when the participants' life values and their paradigm need to be changed first before they achieve the skill given by the training

(I think this is the difference between a training given by a profit company to its staff with the trainings for community development).

I actually had experience about this.

I once conduct trainings on Joyful learning methods in class, which also advocate that physical discipline should not be given to students.

The training participants are teachers and school principals that believe that one effective way to teach discipline to students

is by giving physical discipline actions which creates fear and lead to violence towards children.

These teachers and principals born and grew up in a community that believe physical discipline is common things and it is OK to do it,

and this point of view had been lasting for a long time even from their ancestors.

Now, what I'm trying to say is that they do not realize yet that non physical disciplines can actually be a more effective strategy

in teaching students in class, since it creates a more fun atmosphere for students learning in class ----- because they don't have evidence yet.

They need to practice this new knowledge first, find the proof, then the behavior change, and I say it will take a long time, even if it is supported by post training activities for the participants such as on the job trainings, small workshops, Re-fresher trainings, and so on until the desired outcomes achieved.

Perhaps my conclusion is, a training for increasing participants knowledge & skill without a process of changing one's life values and paradigm,

the outcomes will likely easily to be achieved, rather than the training which require life values and paradigm changes.. the outcomes will likely harder and longer time to be

achieved, and from this we can determine when is the suitable period for us to conduct an evaluation towards training effectiveness.

Laura

Dear Sana,

<it is important to stick to the definition of outcome provided by your employer to avoid unproductive

> debate... >

Yes, I have learned the hard way that the employer or contractor calls the shots, even if he's supposed to be the evaluand. However, arbitrary time dimensions don't change the nature of phenomena. You can call something an impact, or a "surrogate" (proxy) for an impact, but that doesn't mean it is really helping people or the environment, just because you measure it a few years later.

In Brazil in the nineties, putting "institutional development" funding from World Bank or IADB loans into a UNDP (later, also Unesco) project to do "capacity development" often

became an excuse for doing nothing but hiring "consultants" to do the work of bureaucrats, importing computers through UN channels, travelling and providing "training." Sometimes program evaluations done years later could find no evidence whatsoever that Health Ministry projects had improved health (or even attempted to do so), Education Ministry projects had promoted schooling (much less student learning), or Environment Ministry projects had improved (or even protected) the environment. Actually, evaluators were sometimes strongly discouraged from even looking for such impacts.

Suggestion: if an initiative is intended to do nothing but help the learners themselves learn, let's not call it training; let's call it education.

Robert

Dear Vardhani,

Here is my suggestion.

I don't know how frequent the training is being provided to the teachers. But, let assume there are 3 phases of training in one year, and they will be done every 4 months. I'd prefer training's phases/stages rather than a series of training if they are meant to build the capacity of beneficiaries gradually. Say, every phase is a prerequisite for the next and every participant must reach a pre-defined level of practice/capacity (i.e. outcomes) in order to continue to the next phase. In this case, you could assess/evaluate the effectiveness 2-3 months after one phase, or a couple of weeks before the next phase is carried out. Of course you need to carefully define measurable indicators for outcome of each phase – in your case probably are knowledge, attitude and practice/skills. The challenge with this approach is that when the number of participants who goes to the next phase is gradually decreasing in contrast to increasingly difficult to grasp materials. You need to identify immediately (of course through formative M&E) what causes this – whether the design of the training, too high expectation, weak human resources, the way training materials are being provided, etc. – in order to keep the project afloat. While this approach is said to be highly adaptive (theoretically), in practice it is no secret that not easy to change project work plan & direction once it starts. In my opinion, there should be a continuous post-phase accompaniment (mentoring) in order to help those who did not pass a training phase, so in the end all the participants will eventually continue to the next phases and graduated with expected knowledge & skills. But, I am not sure whether this could be done easily as most project-based training often taking assumption that it is designed perfectly and participants do not need to repeat it... :D

Hope my story helps you.

Regards,

Ari...

Hi,

In measuring training effectiveness, another critical dimension to discuss is whether it is a one-off training or a series of trainings for the same group. For eg: in our projects we conduct trainings for school teachers. Each teacher undergoes a series of trainings to ensure that they have adequate capacity to handle a disabled child. These trainings are spread throughout the project lifetime.

In such a case, how does one measure the effectiveness and at what frequency?

Vardhani

robert and others,

i would think that it actually does not matter that kirkpatrick was developed for the business sector and training for business improvement (sales?). i feel that in the non profit sector we should also be concerned with business improvement; that is, if we train people (build their capacity), i assume we do this with the aim that the trainees will do their job better. and eventually we improve the lives of those we work for (poor, children, women, vulnerable groups, etc.).

i have not thought about whether this model will actually be helpful for elementary

education (i doubt, and i think that educators have a host of tools measuring impact of teaching in schools), and would think it is mainly for adult/working people training, of which we do quite a bit at unicef, hence the interest in measuring the impact. several people have also raised the issue of attribution; how can we control for factors beyond the inputs of the training itself (incl. the quality of the facilitator)? but here i would think that at the output level we should certainly be able to claim result (or failure). we are the ones to recruit and select the facilitators, we are the ones that can ensure certain curricula and training circumstances. of course at a higher level than that, attribution plays a role. but the point is that many of us do capacity building as an activity and an expected result/end (very much the human rights based approach to development cooperation)-- so, we should consider evaluation of this intervention as any other project/major intervention. and as such attribution comes in, no escape indeed. but is this not same for any development project?
cheers, dennis

Sampling Methodology

Dear all,

Thanks a heap for the many incisive and practical contributions you have been making on the M&E field.

I am interested in a suitable sampling methodology for a baseline survey for a project using a micro-catchment approach to implementation. The project strives to ensure participatory management of forest and water resources and also support livelihoods improvement of communities living with these resources.

Initially I thought of using shape files of the intervention areas and population data as a basis for estimating sample population and determining survey sample based on the required levels of stratification using the GIS techniques. However, shape files for the intervention areas are not available. The use of administrative boundaries is also not feasible since the intervention areas are marked by physical features and straddles different administrative units.

I will appreciate if the group members may share with me their experiences and alternative approaches to sampling in such a scenario that will equally lead to collection of credible baseline data for assessing progress and evaluating the impact of the project on resource health and the community.

Thanking you,
Shimon Otieno
Kenya

Feasibility study

Dear All,

The NGO which I'm working for is planning to conduct feasibility study for its programme expansion. We are currently working in Tamil Nadu and planning to conduct feasibility study in Andhra Pradesh and Karnataka. Our NGO focuses on Youth development. the main goal of the organisation is to "enable young people to be at the forefront of change and development". the main objectives are 1. to enable young people to take lead role in decision making 2. to enable young people to take responsible decisions regarding Sexual Reproductive Health and 3. to enable young people to have improved life skills opportunity
I would like to know

How do we go about conducting the feasibility study?

which areas should we focus/ with whom should we conduct?

do any one of you have any feasibility study reports to share with us?

what tools should we use?

Regards,
Sylvia

M & E Practitioners' Survey

Dear Colleagues,

I am conducting a study /survey among M & E Experts on the conduct of Midterm Evaluations. I had an initial survey and results but I would like to expand my sample size. I only limit this open type questions for I know you are all busy. I would appreciate very much if you could find time to answer these key questions:

1. What do you usually look into when conducting a Midterm Evaluation of a Project? What documents will you see first?
2. How do you assess the impact of the project at the midterm point?
3. What are the lessons learned you have gained after conducting a project evaluation?

4. What are your recommendations or suggestions to have an effective Midterm Evaluation?

5. What projects & donor agencies you have evaluated?

You can make your answers brief but concise. Please email it back to me in this email address. or answer right below the questions when you reply or separate attached sheets. Thank you very much for your cooperation, God bless you

Sincerely,

Orlen C. Ocleasa

Monitoring & Evaluation Coordinator

ACDI VOCA Philippines

Dear Orlean

First you need to establish if there are: Project documents, Proposal, Logical framework and look at indicators at all levels- but focus on inputs, activities, outputs and outcomes. You then ask efficiency, Effectiveness, relevance, impact and sustainability related questions against the result chain. You can also ask challenges and achievement against the objectives. Check how you are performing against the indicators. Critically discuss with partners/stakeholders to identify lessons learnt after undertaking stock of challenges/obstacles in project/programme implementation. Usually, implementors will list challenges and also identify lessons/ best practices. In conducting Mid term evaluation you are comparing mid-term performance with set targets. You want to know if you will achieve your targets. if you are not achieving the targets then the questions, Why? What can we do, What strategy? What is working and what is not ? etc.

Thanks,

Samuel

M and E officer,

UNICEF-Uganda CO.

Beneficiary Counting

I am interested in examples or experiences with beneficiary counting for multi-sector programming.

I recently started a new M&E position with an international organization with programs in 186 countries ranging from emergency relief to livelihoods. Consequently, the scale and scope of programming and resultant beneficiaries is tremendous, ranging from those who receive a house to condoms, or indirectly benefitting from a public lifeguard, establishment of a disaster early warning system, or construction of a hospital.

I am reminded of a caution from one of my early M&E mentors of the "McDonald's approach," which measures the billions served but says nothing about the quality (nutritional value) of what is being served. However, there is strong interest in being able collect

beneficiary data and report on collective outreach.
Therefore, I welcome any input on beneficiary counting for multi-sector programming.

Respectfully,
Scott Chaplowe
Senior M&E Officer
IFRC

Scott!
Interesting observation "measures the billions served but says nothing about the quality".
Let me ask even a more basic question. Leave aside quality, **what does normal practice actually count?**

Different methods can give different counts. E.g. If an individual opts for 25 days in a Cash for Work program, the individual can be counted once through one system and 25 in yet another system, generating wide variations in count projections.
Unless and until methods for beneficiary counts are fully disclosed, such statistics touted end up as being a little disingenuous.

Rajan Alexander
Development Consultancy Group
43, Da Costa Layout, II Cross,
St. Mary's Town,
Bangalore 560 084

Scott,
dont know if this answers your question but am focusing more on the qualitative.
depending on the sampling size, i wonder if there is a benchmarking conducted prior to the start of the intervention. If it is already in existing, one use participatory M & E, using MSC to capture qualitative impact or even emerging impact of the project. What can be captured may be direct or indirect - I remembered, reviewing a PRA report, on a forest management project. However, as i have reviewed other documents attached (health reports and education report), apparently, there was also the issue of toxicity, as evidenced in the yellowish skin of the children. In my discussion with the community leaders, i said, based on their problem tree, this will be another project impact - other than the usual measurement (e. number of trees, decrease in water turbidity, change behaviour in environment conservation, etc).
While there might be a mandatory indicators, i think using the MSC or even AI, other project impact that is demand driven based on a specific situation will be captured.
ann lily

First off, I want to thank all for the time in your responses. The points raised are extremely relevant, and they reflect the challenges of "beneficiary" counting.

I do like StC's approach of using "reach," which kind of disengages the term from impact/quality.

The question was asked, why count beneficiaries. At this stage, the interest is in reporting on collective outreach. After instilling a culture of collecting and reporting on such output counts, the vision is to scale up to higher level, outcome indicators (KAP -knowledge, attitudes, practice).

Double counting is certainly a challenge, that we are well-aware of from the multiple program setting of the tsunami recovery effort, where there was much overlap among programs and "beneficiaries." Of course, this is complicated by counting direct beneficiaries, i.e. those who have received a shelter, well, or septic system, versus indirect beneficiaries, i.e. those benefitting from the establishment of an early warning system, pre-positioned emergency stocks, or the construction of a hospital. At this stage, I agree that

an important part of the solution is to be clear and consistent on the guidelines for conducting such counts.

Again, thanks to all for the valuable input.

Scott Chaplowe

M&E Senior Advisor

Dear all,

Further to the discussion on the topic Beneficiary count, I would like to share some of our experiences in this regard;

Beneficiary count is an important global activity of Save the Children and we annually count the "Reach" covered by our main 4 Result sectors. We also believe that it is the count of coverage and has nothing to do with impacts or quality of program, for which we adopt different methodologies. I can not go into the details of how we count, however, I would like to highlight some key issues which we encountered and developed some remedial measures:

1) Clear concepts and guidelines who to count, who are direct and who are indirect beneficiaries are very important. These may vary from program to program, and even activities to activities, so this requires clear guidelines according to the program sector and types and nature of beneficiaries. At the top of it, we also need to know why do we need to count? If needed, then to what level of details and how these data will be used?

This will of course depend on Agency priority, importance and need of Beneficiary data.

2) Keeping good data base of beneficiaries reached by the Program sectors is important to correctly count the beneficiaries. Unless you have some kinds of database/templates and people are regularly keeping it update, beneficiary count is a difficult task

3). Duplication issues: Different types of duplication issues exist: duplication created by repeated beneficiaries of same activity, duplication of same people but by different activities, overlapping of multiple programs, duplication between direct and indirect counting and so on. No need to say that if you are not careful, you will be counting the same beneficiaries so many times. We have found that most difficult one is to eliminate cross-sectoral duplications or duplication created by multiple programs in the same geographic areas.

Good knowledge of programs, and nature of beneficiaries, good records of beneficiaries data and agreed methods to eradicate duplication are necessary to avoid the duplication. But note that when we have multi-sector programs, we can only estimate the duplication in many cases, there is no accurate methods as yet. The important precaution you need to take is that, as mentioned below, we do not cross beyond the population in our counting.

4) Concepts and guideline orientation not only to the M&E staff but all program staff who in fact deal with the beneficiaries is most important. Save the Children has a global guideline on Beneficiary reach with concepts and sector-specific guides on how and who to count. If some of you are interested, I can forward.

Many thanks

Dinesh Uprety

Sr. M&E Specialist, Save the Children, Indonesia

Allan!

You write: "I am reminded of coverage boasted by the Marshall Islands Health Department nutrition programs. They counted the number of attendees at their nutrition classes (presumably from attendance records) which exceeded the population of the country." That should not come as a shocker to most of us.

Let say the class strength is 20 and the program is for 30 days. Assume 100% attendance everyday. Assume further daily attendance per class is considered the basis for counting, not number of students passing the course. So for 20 students, the counting could be as high as 600!

It gets even better if these figures are used for further "analysis". Total cost of training lets assume as \$6000. For the count of 600, average cost is a mere \$10 per individual (great "efficiency"!) while for a count of 20, it becomes as high as \$300 (poor "efficiency") I am not sure whether such distortions arise from pure ignorance or deliberate disingenuity practiced. What's sure is that this is sadly very much part of development reporting!

Rajan Alexander
Development Consultancy Group
43, Da Costa Layout, II Cross,
St. Mary's Town,
Bangalore 560 084

How to calculate t-test, chi square or any test!

Dear friends,

Currently I am doing analysis and evaluation report for a 7 months pilot project. Initially baseline data were collected and the same tools with same sample size and study area were also used to collect end evaluation. I able to get some findings by using simple statistics in SPSS. But questioned was raised from some groups to show them the population were similar so that to accept the findings. So pleas help me in this issue. I will be happy if the help is immediate.

Eshetu B.

Hi

It sounds like what you need to do is to get some demographics of each group, that is from the baseline data and then from the end evaluation group. Compare the demographics and see whether they are similar, do they have similar average age, percent male / female, and so on.

Does that help?

Gene Shackman

I agree with Gene; the question you are asking does not need to be evaluated with SPSS. If however, you are asking whether there are any significant differences between the groups, once you have found that they are similar, then you could use SPSS. But it is quite difficult - take it from someone who is currently studying a course on advanced statistics. You could go into analyze/compare means and that will give you access to t-tests, or ANOVA which will give you an F-statistic. Luckily SPSS tells you whether this result is significant or not (if $p < 0.05$). You need to decide which are the independent variables (eg your project inputs) and the dependent variables (the outcomes) and put them in accordingly. Then you should get the result. But SPSS is not for the faint-hearted - and then you would still have to explain the results to your colleagues. Would you feel confident enough to do that?

It sounds to me like the results you got may not have entirely pleased the groups who are asking the questions. Could that also be at the root of the problem?

Pete

What is Black Box in M&E Field

Dear All,

I am Imtiaz and working as M&E Officer.

I am not clear of the term Black Box in the M&E field.

Please guide and if you could furnish with some literature.

Best regards

Imtiaz

M&E Officer

Dear **Hailemichael Taye**

I hope you may have some clarification by reading the beautiful documents available in the weblinks.

http://www.wca-infonet.org/servlet/BinaryDownloaderServlet?filename=1018271895315_SJebellie_Paper8.pdf

<http://www.itrc.org/papers/benchmarkingirrigation/benchmarkingirrigation.pdf>

S.Srinivasan

**Consultant (USAID)- M&E,
Statistics division, MoHFW
Nirman Bhavan, New Delhi**

Hi,

In M&E, I think black box is the area where actions are taken to design M&E system until and unless outputs, outcome, effects, impacts and/or results are not obtained. If we suppose a water reservoir, it usually has an inlet and then water goes through some process and after that it flows out through outlet. This process which starts soon after the input and before it flows out of outlet is a black box. In the project management cycle it carries the same concept behind. So it could be then the planning, implementation strategy, monitoring and evaluation, retreat and policy reform in a project before we get the required result of a project. Similarly, it could be the implementation strategy of M&E such as what tools to be used in M&E, how the tools will be used, who will use it, when will it be used, what data/information we need and from whom, at what size, at what time, what will happen to data, how and who will process it, what will be reported and to whom it will be and how can the report help the project. This is the black box in M&E where stakeholders such as designers, project staff and decision makers need to coordinate and participatorily put together efficient and effective approach and methodology to get the required result from M&E.

Colleagues can correct me if I am wrong 😊

Regards,

Dad

To add on, I have cut and paste from world bank nutrition resources. I remembered in one of the projects I was involved our team had a lingering concern on project evaluation focusing on anthropometric data collection alone during mid-term and end evaluation... Read below.

In Vietnam, community health and nutrition workers observed that, despite comparable socioeconomic status, some children were growing adequately while most suffered from varying degrees of under-nutrition. Using the positive deviance* approach, the behavioral characteristics of "successful" households were identified.

These behaviors included the collecting of shrimp and crabs while working in the fields and supplementing a child's diet with them. An educational campaign promoting such positive deviant behaviors was implemented. After several months, an evaluation was conducted to investigate the effect of the project on nutritional status. Evaluators were pleased to discover that nutritional status had substantially improved since the baseline. However, the evaluation collected only anthropometric data and neglected to determine whether the adoption of positive deviant behaviors had, in fact, taken place.

While the overall improved growth of children in the project area is cause for celebration, the conclusion that this resulted from the educational campaign may have been incorrect. In fact, deworming medication had been introduced into this area during the same period and may have contributed substantially to the decreased rates of under-nutrition.

Simply examining impact relegates the reasons for change to a "black box". Unfortunately, this approach to evaluation is all too common: improvement in impact indicators, where it is observed, is assumed to be attributable to the project without examining process (what we later will define as output and outcome) indicators.

*Positive deviance refers to situations in which individuals or households are doing better than would be expected given their social and/or economic circumstances. Their time and resource allocation strategies may be worth disseminating more broadly.

Ann Lily Marie O. Uvero

Dear Imtiaz,

The term "black box" came from the so-called "black box evaluation model." It is the origin of interventions model that develop under program theory. The model is usually presented as follows.

Inputs -> "Black box" -> Outputs

The black box could be anything, it can be an organization or an intervention. It is assumed that if one give inputs to it, it will produce outputs. In brief, black box is something that will "change" inputs into outputs. In M&E, practitioners often equal black box with processes or activities. Also, it is considered that black box model is contextual in nature in which external environment (e.g. social, political, economy, environment, etc.) could affect inputs, outputs and the black box itself. It should be emphasised here that this is the primitive model of evaluation.

From this concept, then the logic model as follows developed afterward.

Inputs -> Intervention: Process/Activity -> Outputs -> Outcomes -> Impacts

I suggest you to read Weisz (1996) classic book of evaluation to learn more on this black box model, or other text on program theory.

Hope this story helps you.

Regards,

Ari...

Dear colleagues

My understanding, and use, of the term is that a "black box" approach is one which reports on outputs or outcomes without any any analysis of the underlying causal mechanisms (program theory, logic etc) by which the program inputs have resulted in affecting those outputs/outcomes. Lest there be any doubt, this is a bad thing!

So in the area of impact evaluation, approaches relying solely on regression-based estimates from the coefficient on the project dummy variables, or control group-based estimates simply looking at mean differences in outcomes between treatment and control, are often guilty of adopting a black box approach. The draw back of this approach is that it misses possible policy lessons that come from unpacking the causal chain.

As an example, I was involved in the evaluation of a nutrition project in Bangladesh which found little impact on nutritional status, though allowing for impact heterogeneity (looking at how impact varies across groups) found that there was an impact on the most malnourished. The most malnourished were of course the target group, but the data showed substantial mis-targeting, with many of those who should benefit not doing so (Type I error), whilst many in the program should not have been there (Type II error). Why was this? Admission to the program was through growth monitoring, in which community workers weighed each child monthly and mapped the child's progress on a growth chart. It turned out that these community workers could not correctly read the growth charts to identify which children ought be admitted to the program. Hence a clear reason for low impact (children who will not benefit are getting the intervention), and a policy implication (improve the quality or training of community workers). (There were other reasons, I have picked one for illustrative purposes).

There is a brief discussion of black boxes in my pamphlet on impact

evaluation: <http://ideas.repec.org/p/pramprapa/1111.html>
and the Bangladesh project is written up in
<http://ideas.repec.org/a/wly/jintdv/v19y2007i5p627-652.html> and
<http://ideas.repec.org/p/wpa/wuwpdc/0510004.html>

Howard White

www.3ieimpact.org

www.impactevaluation2009.org

Impact Evaluation of a Mass Media Organization

Dear all

Presently I am involved in 5 year impact evaluation of a Mass media (TV documentary, radio, mobile broadcasting unit) organization in Cambodia, with focus on women issue (Domestic violence, women rights etc).

They do not have a baseline indicator. How to measure the change. Please share any resources on similar evaluation.

Thanks in advance,

With my best regards

Umakant Singh

lecturer and freelance consultant

Hi,

Based on my experience, when the mandate is to measure the change after a developmental intervention in situations where baseline data is not available, deeper research is required through open ended questions especially through focus group discussions etc on the condition of any specific area where improvement has happened as a result of an intervention. There should be specific examples and information about past conditions. Researcher needs high facilitation skills to grip and grasp information from interviewees. It could include qualitative and quantitative information. In order to measure the change, we would also need to measure improvements and the immediate effects and/or impacts of the developmental intervention at the same time. During analysis of the data, we might be able to compare past and present conditions of a given situation and will be able to measure the change.

Regards,

Dad

Dear Umakant,

What I would suggest you do is the following as you don't have access to baseline data::

1. Conduct a focus group among the key individuals who have been involved and know about the situation of the program from the beginning of the program and even before the program had actually started its implementation. Prepare a questionnaire that focuses on the situation of the subject matter before your program interventions were delivered. They would have ideas about the situation in the past and might answer to your questions with their good information.
2. Interview the key stakeholders of the program separately and collect their ideas about the program and particularly what the situation was before the program was implemented. Later you can consolidate their ideas and information and then consider that information as the baseline.
3. You might have access to secondary data that addressess the issues covered in your program both during and implementation of the program and before it was implemented. I think this might help you find at least some ideas on what the situation was before the program was actually implemented and you could consider those ideas and information as your baseline so you could compare it with the changes that have been made to the environment of course with the contribution of your program's interventions.

I hope that helps to some extend...

Best regards,

Hameed Ah. HAQPARWAR

National Consultant - M&E
United Nations Educational Scientific and Cultural Organization (UNESCO)
Enhancement of Literacy in Afghanistan (ELA)