



LEAP

***Learning through Evaluation
with Accountability & Planning***

Evaluation Terms of Reference Guidelines

1st Edition

World Vision International
LEAP Team

LEAP Evaluation Terms of Reference Guidelines
 Published June 8, 2007
 © World Vision International, LEAP Team
Send comments or questions to:
 development_resources_team@wvi.org

Contents

	Page
Acknowledgements	2
Affirmation	2
Glossary	2
Introduction	3
Evaluation Terms of Reference Guidelines	3
Appendices	9
A. Evaluation Purposes Table	10
B. Guidelines for Conceptualisation of an Evaluation	11
C. Evaluation Matrix Example	22
D. Types of Data Collection	23
E. Team Members and Roles Example Table	26
F. Evaluation Action Plan Matrix	27

Acknowledgements

The following parties contributed to the 1st Edition of this document:

The Support Office LEAP Consortium, especially Gisela Poole (WV Germany), Andrew Newmarch (WV Australia), Ari Uotila (WV Canada)

World Vision National Offices whose DME work contributed to this document, especially WV Armenia, WV Vietnam, WV Cambodia, and WV Albania

The LEAP Strategy and Working Group, especially Frank Cookingham

The extended Development Resources Team, especially Ratu Saha

The Development Resources Team

Affirmation

Except as acknowledged by the references in this paper to other authors and publications, the template and guidelines described herein consists of our own work, undertaken to improve the quality of World Vision's Design, Monitoring and Evaluation Learning System.

Lydia Ledgerwood-Eberlien, Coordinator – LEAP Programme Resources
 June 8, 2007

Glossary

ADP	Area Development Programme
DME	Design, Monitoring, and Evaluation
EC	European Commission

HEA	Humanitarian and Emergency Affairs
IFAD	International Fund for Agricultural Development
LEAP	Learning through Evaluation with Accountability and Planning
LNGO	Local Non Government Organisation
M&E	Monitoring and Evaluation
NGO	Non Government Organisation
NO	National Office
PRA	Participatory Rural Appraisal
SO	Support Office
SWOT	Strengths, Weaknesses, Opportunities, and Threats
TD	Transformational Development
TDI	Transformational Development Indicators
TOR	Terms of Reference
UNFPA	United Nations Population Fund
USAID	United States Agency for International Development
WV	World Vision
WVI	World Vision International

Introduction

These guidelines and their accompanying template are just the beginning of the LEAP aligned resources that will become available over the next few years. The content is an attempt to strike a balance between the amount of information needed to properly evaluate a programme with constituent projects and the complex and varied contexts that World Vision works in. The documents are meant to be used in any programme situation, including TD, HEA, and Policy and Advocacy.

Extensive consultation and literature review were conducted both within and without the World Vision partnership to arrive at these versions. In the spirit of LEAP, these documents are in no way final, but will continue to be changed and improved upon as the Development Resources Team receives feedback from practical field use. Please do not hesitate to provide comments to the DeRT to this end.

Please note that the Evaluation TOR Guidelines (and the accompanying Template) are addendums to LEAP 2nd Edition and **do not stand alone**. They must be read as a set to fully understand the what, why, and how of the evaluation process. Likewise, the Evaluation TOR Template should not be filled out without carefully reading these Guidelines.

Evaluation Terms of Reference Guidelines¹

Cover Page

The cover page should be simple and direct. It should identify the document as an Evaluation TOR, including the name (which should indicate whether the programme is an ADP, sectorally based, or other) and number of the proposed development programme, date and phase number of programme (e.g. 2008-2012, Phase II) and National Office. The WV logo should be located in the top right hand corner and meet with agreed partnership branding strategies. (See the cover page of this document for an example).

¹ Adapted from WVI LEAP version 1.0; DME manual_Armenia; HEA Terms of Reference; EC_ToR_suggestions; ToR_Lezha ADP_Albania

Additional information to the above should only be added where justified and provides improved clarity on the context of the design. The cover should not include photos and/or graphics (e.g. borders).

All design documents should have a copyright notice that includes the date the design was produced, along with the copyright symbol e.g., Produced December 2002, © World Vision International.

i. Table of Contents

A list of the main sections of the design document should be presented in this section. It may be useful to add in various sub-headings, marking significant places to find different themes and discussions. It is strongly suggested that the automatic table of contents function be used where possible, as it automatically updates when the document is reformatted from letterhead to A4 or vice versa.

ii. Acknowledgements

Include an acknowledgements page to list the people that have contributed to preparing the design. This will include the principle authors, contributing partners (either people or organisations) and should also include key members of the community as appropriate. Also use this opportunity to thank people who have assisted throughout the process of assessment and design. This could include such people as staff from the National Office and other Partnership offices who have been involved and people from the community who have played a major role. See the acknowledgements of this document for an example.

iii. Affirmation

The affirmation states the motive and objectives of the design that is being presented and also that the material is original work. It would also be appropriate to acknowledge that the intellectual properties of the design rest with the communities about which the design is written.

“Except as acknowledged by the references in this paper to other authors and publications, the design (and assessment research) described herein consists of our own work, undertaken to secure funding, implement the activities, describe and advance learning, as part of the requirements of World Vision’s Design, Monitoring and Evaluation Learning System.”

Primary quantitative and qualitative data collected throughout the assessment and design process remain the property of the communities and families described in this document. Information and data must be used only with their consent.

If this is a re-design, acknowledge the date this design comes into effect and that it supersedes previous design documents.

[insert name of principle authors here]

[insert date here]

iv. Glossary

The glossary is an alphabetical list of terms or words that are found in the document or related to the text of the document that need some explanation or which may help the reader to a greater understanding. The list can expand on the complete terms that maybe acronyms and abbreviations as well as explain the concept of an ADP (Area Development Programme).

ADP	Area Development Programme
RC	Registered Children

v. Introduction

The document's introduction could include the following information:

- How it is the fruit of collaboration/partnership with communities
- Reference to the design document and previous management reports
- Timeframe the evaluation covers and lead up efforts to get here
- Description of the programme's defining characteristics e.g. location, disaster/emergency based, advocacy based, transformational development, high prevalence HIV/AIDS, etc

I. Evaluation Summary

Programme/Project:	Title and number.
Programme Phase:	Period being covered by the evaluation
Evaluation Type:	e.g. interim evaluation, End of Programme or project, post-programme
Evaluation Purpose:	e.g. <ul style="list-style-type: none"> • provide information on what worked, what did not, and why; • determine whether underlying programme and project theories and assumptions were valid; • determine the efficiency, consistency, effectiveness, relevance, and sustainability of projects and programmes; • guide to decision makers or programmers in reproducing programmes that succeed; • encourage and celebrate the achievements of partners; • promote accountability and learning; • See Appendix A: Evaluation Purposes Table, for further possibilities
Primary Methodology	E.g.: <ul style="list-style-type: none"> • Household survey • Participatory exercises with community representatives • Expert review
Evaluation Start and end dates:	
Anticipated Evaluation Report release date:	

- How was this document prepared?
- Who participated in the preparation?

2. Description of Programme or Project Being Evaluated

This section should briefly frame the evaluation context. Do not go into great detail, and reference the design document and appropriate semi annual and annual reports where necessary.

- What project or programme is being evaluated? Outline the context (such as political, cultural, socio-economic, religious situation) and characteristics of the project or programme (including changes made as a result of redesign or reflection exercises, if applicable).
- What are the main programme or project objectives? Any significant changes to the original objectives or plans? If it is an evaluation of an ongoing activity, describe the current state of implementation, indicating any notable successes or problems.
- How long has the programme/project been operational?

- What is the annual budget, number of beneficiaries, major donors, project duration, etc? For programmes: anticipated life-of-programme budget, number and kinds of projects implemented so far and a description of direct (project level) and indirect (people groups, institutions, organisations, beneficiaries) partners.

3. Evaluation Partners (i.e. who is the evaluation for?)

- Who are the groups and individuals who have a major interest in the evaluation? (for whom it is primarily intended, the target audience)
- E.g. communities, partner organisations, national office management staff, Support Office staff, donors, etc.

4. Evaluation Type

- What type of evaluation is this? E.g. Interim, end of programme or project, post-programme?
- What are the reasons for scheduling this type of evaluation?

See Appendix B: Guidelines for Conceptualising an Evaluation, for assistance/detail on this section.

5. Evaluation Purpose and Objectives

- State the Evaluation *Purpose*: See Appendix A: Evaluation Purposes Table for a detailed list of possibilities. Please note that an evaluation should have one and only one purpose or focus. Trying to achieve several purposes with limited resources usually dilutes the usefulness of the evaluation.
- What are the specific *objectives* (including the role of the evaluation team to facilitate learning, how the results of the evaluation will be used, equity issues in terms of which groups will be involved and empowered by the process (gender, poorest of poor, the disabled, etc)) to be accomplished in order to achieve the purpose of the evaluation? What is the rationale for choosing these objectives? Consider:
 - **relevance** What value does/did the programme/project add to its context or situation? How appropriate is/was the programme/project to the situation in the community? How significant and worthwhile is/was the programme/project to the situation. Did we do the right thing?
 - **effectiveness** Have the planned outputs led to the achievement of the outcomes? Have unexpected outcomes happened? Have assumptions affected project achievements? How effective was the risk management? Did any re-design occur and if so why? Did we do the right thing in the right way? This should include a specific assessment of the benefits accruing to target groups, including women, men, children, the elderly and people with disabilities, etc.
 - **impact** The effect of the project on its wider environment, the long-term social change at the community level the project has contributed to
 - **efficiency** Have project outputs been achieved at reasonable cost? Has the project been cost-effective or would other approaches have led to the achieving of the same results at more reasonable costs? Did we do the right thing for the right cost?
 - **sustainability** What is the ability of target communities to maintain the higher level of livelihood standards that have been obtained through our programme/project intervention? Important areas: institutional, financial, economic, technological, socio-cultural, ecological and political. See the Sustainability and Transition Guidance Notes and Tools, which can be found on the TD website, for details. **All programme evaluations (especially the first evaluation in the life of a programme) should include a study of sustainability as it relates to the sustainability table in the programme design document.**

See Appendix B, Table 3: Framework for Prioritization of Evaluation Objectives for further explanation and detail.

- One objective in every programme or project evaluation should be to describe how the programme/project has been consistent or inconsistent with World Vision’s integrated focus for programming. If this objective must be excluded from an evaluation due to, for instance, lack of capacity or resources, note this in the terms of reference and explain in detail why it has been excluded.

- What are the *key questions* being asked? What information is needed to answer these questions, and from where can it be collected? (an Evaluation Matrix, found in Appendix C, can be used to provide more details on key questions and how they can be answered) These questions should be guided by the Monitoring and Evaluation plan outcome and goal level indicators.
- What is the *scope* of the evaluation? Is it the correct size, scale, or scope given the size, scale, and scope of the project/programme being evaluated? Consider the capacity of staff involved and resources available.
- Which specific issues require *recommendations* from the evaluation team? (this could be some or all issues identified).

See Appendix B: Guidelines for Conceptualizing an Evaluation for assistance in developing this section.

6. Evaluation Methodology

- For each of the general three phases of the evaluation (initial planning, data collection and analysis, findings/recommendations and follow-up), explain the methodologies that will be used to gather data. Indicate preferred/possible quantitative and qualitative evaluation techniques and research methods – see Appendix D: Types of Data Collection, for examples
- What disaggregation requirements are there (gender, ethnic group, income level, etc)?
- How much data should be collected (sample size, number of communities, etc)?
- How will the data be analysed? (e.g. what software, what process). Who will “own” the raw data?
- How will the communities, resource people, and other partners be included in the data collection and analysis (including formulation of implications and recommendations), and how will the findings and recommendations be fed back to them?

7. Limitations

- What conditions (such as funding, time, logistics, politics, etc.) limit the evaluation in any significant way?

8. Authority and Responsibility

a. Team Members and Roles

See Appendix E for an example table of team members and roles.

- Who will be on the evaluation team, and what will be their role or area of expertise from the perspective of management and technical skills?²

² Evaluation team members need not be external, particularly for ADP evaluations. Additionally, if we are seeking to do participatory evaluations (as is stated in LEAP 2nd Edition), there should be a mix of staff, community, and external members on the evaluation team. An exception would be if the evaluation is one required by a donor. In that case there may be a requirement that the evaluation team members are external to the programme that is evaluated.

- Indicate the *number of experts* required, and their key qualifications and experience, with special attention to the requirements for the position of **Team Leader**.
- It is also especially important to ensure that, among the various expertise necessary, adequate **gender expertise** is present, members of the **sponsorship team** are included, and **national experts** should be included wherever possible.
- It is vital that at least one team member has good experience of conducting evaluations, ideally the Team Leader.
- If there is a consultant employed, what will be his/her exact role?
- Who will analyse data?

b. Partners (i.e. who will be involved in the evaluation?)

- Outline of the roles of the main partners (e.g. NO, SO, ADP, local government, community institutions) with respect to:
 - Hiring of consultants and/or secondment of staff
 - Provision of requisite documents for literature review
 - Approvals
 - Provision of services, logistics, equipment, supplies
 - Facilitation of meetings and data surveys in the community

It would be useful to present this information in table form.

See Appendix B, Table 2: Framework for Determining Community Participation in Evaluation, for suggestions for organising this information.

9. Team Advisors

- Will there be any advisors participating as either full-time members or otherwise as their responsibilities permit? Who are they and how will they be consulted?

10. Time frame

- Outline a schedule for each event, its duration, and the approximate number of participants. Divide the schedule by the three general phases (initial planning, data collection and analysis, findings/recommendations and follow-up). Include: time for preparatory work (document review, etc), expected dates for the hiring of the evaluation consultant (if applicable), formation of the evaluation team, finalising the evaluation workplan, field data collection and validation of draft findings, data analysis, draft evaluation report, debriefings with key stakeholders (e.g. community, local government, donors, national office, etc), final evaluation report).
- Be sure to allow time for processing and reflecting on data collected at reasonable intervals as well as travel times in-country.

11. Logistics

- Briefly outline how arrangements will be made in-country for:
 - accommodation and transport;
 - locations for the debriefings and the review of the first draft of the report (while the evaluation team is still in country);
 - necessary services such as translators, interpreters, drivers, data processors, facilitators, access to desk space and computers, printers for non-programme evaluation team members.
- If an outside lead evaluator will be utilised, be sure to provide a point person on the Field Management Team to arrange logistical details before and during the evaluation.

It would be useful to present this information in table form.

12. Products

- What products, including the report, will result from this evaluation? How will they be prepared, who will they be delivered to, and by when?
- How will the results/findings be recorded?
- How will findings (conclusions and recommendations) be presented to different stakeholders?
- How will feedback about the evaluation process and results be incorporated into the final results?

13. Budget

- What resources are required for implementation? (Include budget breakdown).
- Provide a short budget narrative if needed.

Note: Please use the LEAP Budget tool template for preparing the Evaluation TOR budget. This is a generic tool that works with any kind of project budget. It can be found on the TD website (www.transformational-development.org)

14. Documents

- List major documents that should be reviewed in the research phase of the evaluation.

15. Lessons Learned

- Describe the process that will be followed to identify and record lessons learned about planning and doing evaluation

16. Appendices

Appendices for the Evaluation TOR should include:

- Evaluation TOR budget
- Tables (team members and roles, timelines, etc)

Appendices to the Evaluation TOR Guidelines (this document) are as follows:

- A: Evaluation Purposes Table
- B: Guidelines for Conceptualizing an Evaluation
- C: Evaluation Matrix Example
- D: Types of Data Collection
- E: Team Members and Roles example Table

Appendix A: Evaluation Purposes Table

Appropriateness of Design and Monitoring systems	To determine how well the program was designed and monitored, against the objectives, benchmarks and processes established in the program design document.
Target versus actual	Assess the progress made towards achieving each Project Goal (or technical sector) based on the current logframe, design and monitoring data.
Impact	Assess progress made with regards to WVI's five Domains of Change within Transformational Development. Determine the impact and/or potential impact of the program.
Strengths/ Weaknesses	Determine strengths (including successful innovations and promising practices) and weaknesses (factors impeding progress) of the program planning, design, implementation M&E, and ongoing community management in the program.
Efficiency and Effectiveness	To determine whether the resources (financial, human & materials) have been used economically and wisely for the well-being of the community.
Equity	To determine how the program involved and benefited different genders, ethnic minorities, children and the disabled throughout the planning, design, implementation, monitoring and evaluation and ongoing management of the program.
Capacity	Assess ADP structure and staff capacity to undertake the development initiatives. Assess the status of stakeholders (primarily HFs and District/Commune PMBs) to design, implement, monitor and evaluate activities.
Ownership and Sustainability	To assess the overall management and structure of the program, particularly focusing on the potential for the program to be owned by local people and for benefits to continue after the ADP may close.
Participation	Assess the level and quality of participation by women, children, people with disabilities, and the poor.
Recommendations	Provide specific, actionable, and practical recommendations for focus on key sectors during the next phase.

Appendix B: Guidelines for Conceptualising an Evaluation

There are a number of factors or variables that will determine the course an evaluation will take. It is important that those responsible for early evaluation planning carefully consider a number of questions/criteria, which when analyzed, will facilitate the setting of the evaluation purpose, evaluation objectives and will guide subsequent evaluation questions and results. The intent of this section is to introduce some of those high level questions/criteria and provide some tools and metrics to help the evaluation planner to set the overall evaluation purpose and objectives as well as provide guidance on subsequent evaluation questions.

Considerations for conceptualizing the evaluation

I. Independent or Participatory Evaluation Approach.

One early consideration in an evaluation process is to consider whether or not an evaluation needs to be an “independent” evaluation or a “participatory” evaluation. While most evaluations will be neither fully independent nor fully participatory, this matter is an important consideration in evaluation design. Typically an “independent” or “external” evaluation is one where:

- The evaluation terms of reference may have been to a large extent set by an agent external to the programme or project
- The lead evaluator is external to World Vision
- Possibly all the members of the evaluation team are people external to the programme/project.

Such external evaluations are often required by major/governmental donors to comply with funding requirements. Sponsorship-funded programmes (ADPs) and projects tend to use more participatory and community-based evaluation approaches, as proscribed in LEAP 2nd Edition. Table I below identifies some of the advantages and disadvantages of using an “external” or “internal” lead evaluator.

Table I: Trade-offs Between Internal and External Evaluators

Lead Evaluator associated with the project/programme (Internal)	
Advantages	Disadvantages
<ul style="list-style-type: none"> • Knows World Vision and/or the programme/project. • Understands and can interpret behaviour and attitudes of members of the organization. • May possess important informal information. • Is known to staff, so may pose less threat of anxiety or disruption. • Can more easily accept and promote use of evaluation results. • Is often less costly. • Doesn't require time-consuming recruitment negotiations. • Contributes to strengthening national evaluation capability. 	<ul style="list-style-type: none"> • May lack objectivity and thus reduce the external credibility of findings. • Tends to accept the position of the organization – may limit exploration of alternative solutions • Is usually too busy to participate fully. • May be part of the authority structure and may be constrained by his/her organizational role. • May not be sufficiently knowledgeable or experienced to design and implement an evaluation. • May not have special subject matter expertise.

Lead Evaluator not associated with the project/programme (External)	
Advantages	Disadvantages
<ul style="list-style-type: none"> • May be more objective and find it easier to formulate recommendations. • May be free from organizational bias. • May offer new perspective and additional insights. • May have greater evaluation skills and expertise in conducting an evaluation. • May provide greater technical expertise. • Able to dedicate him/herself full time to the evaluation. • May be able to serve as an arbitrator or facilitator between parties of stakeholders. • May be able bring WV/programme/project into contact with additional technical resources. 	<ul style="list-style-type: none"> • May not know World Vision and its values, policies, procedures, and personalities. • May be ignorant of constraints affecting feasibility of recommendations. • May be unfamiliar with the local political, cultural and economic environment. • May tend to produce overly theoretical evaluation results (if an academic institution is contracted). • May be perceived as an adversary arousing unnecessary anxiety. • May be costly. • Requires more time for contract negotiations, orientation, and monitoring.

Adapted from UNFPA Programme Manager's Planning, Monitoring & Evaluation Toolkit, 2004.

To determine (either before the evaluation or afterwards) the level of community participation in an evaluation, a number of criteria should be considered in addition to the qualifications of the evaluation team leader. Attached below (Table 2) is a framework that could be used to determine the level of community participation in an evaluation. As appropriate, this information could also be incorporated into the evaluation terms of reference in order to dialogue with the donor/donor's representative.

Table 2: Framework for Determining Community Participation in an Evaluation

Task or Decision		Who will make the decisions? Who will do the work?			
		External Evaluator	Program Staff	Program Participants	Other Stakeholders
A	Decisions about what information is to be collected.	1	2	3	3
B	Decisions about how to collect information.	1	2	3	3
C	Collection of information.	1	2	3	3
D	Organization of information (tabulation).	1	2	3	3
E	Decisions about what the information means.	1	2	3	3
F	Statement of conclusions.	1	2	3	3
G	Statement of recommendations.	1	2	4	3
H	Writing the report draft.	1	2	3	3
I	Comments on the draft that will lead to appropriate revisions.	1	2	3	3
J	Control over who receives the information in the report.	1	2	3	3
K	Who can understand the written report?	1	2	3	3
	TOTAL				

One possible interpretation of scores

00 – 15 = external evaluation > 'Participatory Score' = 1
 16 - 35 = level I participation > 'Participatory Score' = 2
 36 - 70 = level II participation > 'Participatory Score' = 3
 71 - 100 = level III participation > 'Participatory Score' = 4

Source: Frank Cookingham. 2002. "Participatory Processes in the Community." WVI.

2. Evaluation Type

In broad terms, evaluations can be classified into two types: formative and summative. Formative evaluations are usually taken earlier in the programme/project cycle or lifetime in order to gain a better understanding of what is being achieved and to identify how the programme/project can be improved.

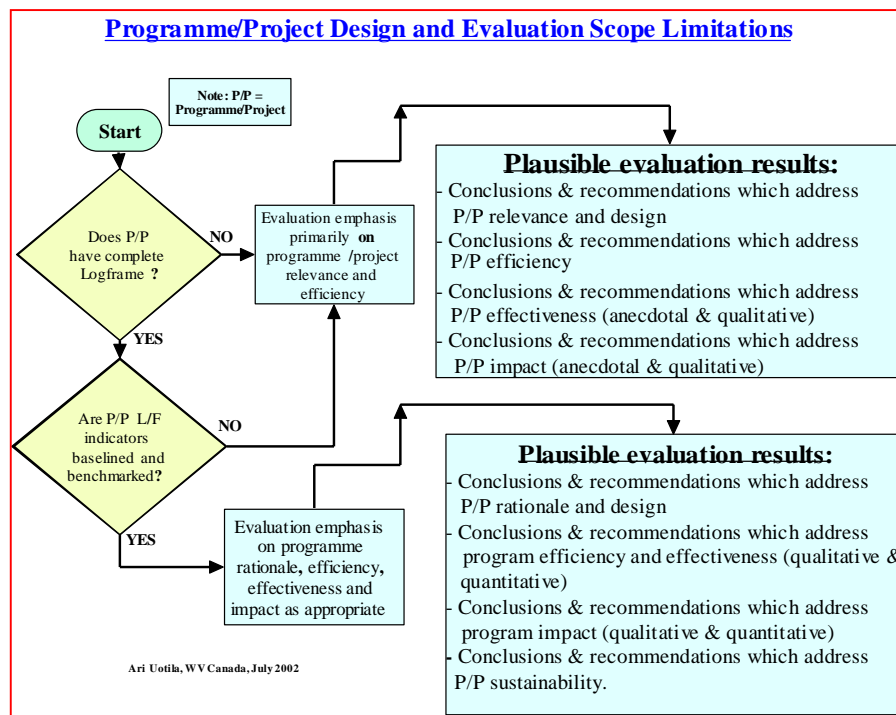
For the purposes of LEAP, formative evaluation is covered under Assessment. Summative evaluations are usually undertaken later on in the programme/project life in order to establish the effectiveness and value (worth) of the programme/project. Summative evaluations focus more on the long-term outcome and goal-level results of the programme/project. Thus, the type of summative evaluation to be determined in an ADP evaluation would be either:

- Interim evaluation;
- End of programme/project evaluation; or
- Post programme evaluation (also called *Ex Post*)

3. Quality and Completeness of Programme/Project Design.

The quality and completeness of the design of the programme/project has a major bearing on what can subsequently be evaluated or not. The two most common issues in practice are poor/incomplete indicators in the Logframe, and lack of or incomplete baseline data and benchmarks for those indicators. Without indicators and benchmarks, it a) becomes very difficult to monitor normative progress of the programme/project, b) is almost impossible to speak to a programme/project's contribution to changes in the lives of community members, or c) is difficult to attribute changes seen in the community to the project/programme. Additionally, without a sound logframe, the articulation of programme theory will be insufficient to use a "programme theory-driven" approach to evaluation. Figure 1 below identifies some of the constraints on evaluation imposed by poor/incomplete programme/project design.

Figure 1: Quality and Completeness of Programme/Project Design and Evaluation Scope



Adapted from "ADP Design and How it Affects Monitoring and Evaluation." Ari Uotila. WVC. 2002.

4. Evaluation Purpose

See Appendix A for examples of evaluation purposes.

5. Evaluation Objectives

As discussed elsewhere in this section, the main evaluation objectives (foci) categories are **relevance, efficiency, effectiveness, impact and sustainability**. The evaluation objectives are a key precursor in determining the broad evaluation questions, which in turn will guide the formulation of the subsequent evaluation data collection methods. How does one determine the evaluation objectives? The foregoing criteria (independent or participatory, formative or summative, quality and completeness of programme/project design) will certainly impact the evaluation objectives. Other factors that will also influence evaluation objectives include:




- age of project/programme
- variability/changeability of the programming environment
- programme/project management environment
- donor special needs/requirements
- programmatic risk

Below is a framework (Table 3) that suggests means and ways to prioritize evaluation objectives using the criteria discussed above. It is important to keep evaluations as focused as possible in order to keep down costs, minimize organizational and community disruptions, and in order to produce the evaluation results in a timely and efficient manner.

Table 3: Framework for Prioritisation of Evaluation Objectives

Relevance		
Considerations/Questions to ask	High/Low Priority	Indications
<p>What value does/did the programme/project add to its context or situation? How appropriate is/was the programme/project to the situation in the community? How significant and worthwhile was/is the programme/project to the situation. Did we do the right thing? Are there changes in the environmental, social, political, security or economic situation of the programme/project, which necessitates reformulation of design?</p> <p>Alternative Strategies: Is there evidence of better (alternative) strategies or ways of addressing the identified problems/needs of the community? How can these be incorporated into a revised design or redesign?</p>	High (Score 3) ↑	<ul style="list-style-type: none"> • High risk/changeable environment • “Young’ project/programme • Programme/project operating environment has changed • Poor/incomplete programme/project design • Problems in relationships/management
	Medium (2)	
	Low (1) ↓	<ul style="list-style-type: none"> • Stable environment • Good management • Little change in overall programming environment • Good & complete programme/project design
Efficiency		
Considerations/Questions to ask	High/Low Priority	Indications
<p>Have project outputs been achieved at reasonable cost? Has the project been cost-effective or would other approaches have led to the achieving of the same results at more reasonable costs? Did we do the right thing for the right cost ? Are there more efficient ways and means of delivering more and/or better outputs with available inputs?</p>	High (3) ↑	<ul style="list-style-type: none"> • Need to quantify project/intervention costs per beneficiary • Mid-term project assessment • Need to compare alternative methods of intervention delivery • Good M&E plan and tracking of financial and economic (inc. community) inputs
	Medium (2)	
	Low (1) ↓	<ul style="list-style-type: none"> • Efficiency assessment remains valid and needful at all stages of the programme/project. Monitored primarily at the project output level.

Effectiveness		
Considerations/Questions to ask	High/Low Priority	Indications
<p>Have the planned outputs led to the achievement of the outcomes? Have unexpected outputs happened? Have some of the design assumptions not held and if so, how has this affected project/programme achievements? How effective was the risk management? Did any redesign occur and if so why? Did we do the right thing in the right way? This should include a specific assessment of the benefits accruing to target groups, including women, men, children, the elderly and disabled etc.</p>	High (3) ↑	<ul style="list-style-type: none"> • Evaluation needs to show/speak to causality of results (attribution/ contribution to changes) • End of project or end of programme cycle evaluation • Good and complete design and baselines • Need/ability to show utilization of project outputs at community level
	Medium (2)	
	Low (1) ↓	<ul style="list-style-type: none"> • Newer project/programme • Short term project intervention • Poor/incomplete design or baselines.
Impact		
Considerations/Questions to ask	High/Low Priority	Indications
<p>The effect of the project on its wider environment, the long-term social change at the community level the project has contributed to.</p> <p>Causality: Where causality needs to be demonstrated or addressed by the evaluation, two methodological aspects need to be considered: firstly, whether to evidence causality through <u>attribution</u>. That is changes in the community observed are attributable to programme /project (typically statistical significance needs to be demonstrated).</p>	High (3) ↑	<ul style="list-style-type: none"> • Evaluation needs to show/speak to causality of results (attribution/ contribution to changes) • End of programme cycle, end of programme evaluation or ex-post evaluation. • Good and complete design and baselines • A quasi-experimental evaluation design if attribution is to be addressed
	Medium (2)	

<p>Secondly, whether the programme/project has contributed in meaningful and plausible ways to the changes observed in the lives of community members. This is more of a qualitative and a “reasoned argument” approach and does not require the statistical rigour demanded by evidencing of attribution.</p> <p>Key considerations with respect to causality and programme/project impacts include: an examination of particular factors or events, which may have affected results; and a determination needs to be made whether these factors are internal or external to the project/programme. It is the “internal” factors (included in project design of in the programme theory), that provide the basis of project contribution or attribution arguments.</p>	<p style="text-align: center;">Low (1)</p> 	<ul style="list-style-type: none"> • Project evaluation • Newer programme or a mid-term evaluation • Poor/incomplete design • Causality does not need to be addressed by evaluation
<p>Sustainability</p>		
<p>Considerations/Questions to ask</p>	<p>High/Low Priority</p>	<p>Indications</p>
<p>What is the ability and readiness of target communities to maintain the higher level of livelihood standards that have been obtained through our programme/project intervention? Assess ability and readiness to replicate interventions in non-intervention communities.</p> <p>Assessment of availability of local management, institutional, financial, economic, technical and socio-cultural/political resources needed to maintain the programme results after withdrawal of WV resources.</p> <p>Assessment of Programme Transition Strategy needs to take place and would include an assessment of progress towards transition benchmarks using the transition indicators, namely:</p>	<p style="text-align: center;">High (3)</p> 	<ul style="list-style-type: none"> • End of programme cycle, end of programme evaluation or ex-post evaluation. • Better suited to a participatory evaluation approach • Good and complete design including complete programme sustainability and transition strategies and associated project sustainability and transition plans
<ul style="list-style-type: none"> - Progress in quality of life indicators, mainly related to well being of children (This would include selected TDIs) - Community capacity indicators - Transfer of responsibility indicators 	<p style="text-align: center;">Medium (2)</p> <p style="text-align: center;">Low (1)</p> 	<ul style="list-style-type: none"> • There is never a “low” need for a sustainability focus in programme/project evaluations – but this assessment is periodic in some respects (evaluation based) – but also needs to permeate programme monitoring and the programme ethos.

Demonstration of the use of the Framework for Prioritisation of Evaluation Objectives

Not all of the five main evaluation objectives should receive equal prioritisation in any one evaluation. The number and scope of the evaluation question outlined in the first column “Considerations/Questions to ask” will be determined by the “Indications” and the assessed “value” from 1 to 3 that is given each objective. While there needs to be a caution attached to the use of the numbers in the “High/Low Priority” column, the use of the numbers maybe helpful in outlining the evaluation objectives graphically.

Let us consider the following illustration:

ADP Azure is ten years old and needs to have it’s second evaluation. The first evaluation was conducted five years ago, and while the evaluation report was completed, it did not result in any significant redesign due to a change of ADP management and the lack of timely response from the Support Office on the evaluation report. The original design of the ADP contained a good logframe with good indicators, but no baseline survey was conducted. TDI indicators were measured in the ADP three years ago, and now the plan is to have the ADP become “LEAP-aligned” in terms of its new (re) design. The ADP is funded entirely through sponsorship funding except for a food security project, which is part of a country-wide food security programme funded by Government “X” through Support Office “Y.” The food security programme evaluation requires an external evaluation team leader, and as it started two years ago, it will be due for an evaluation in one year’s time. The ADP area experienced a major drought three years ago, which caused 50% crop losses and 20% livestock losses in one year. The situation has been more stable subsequently.

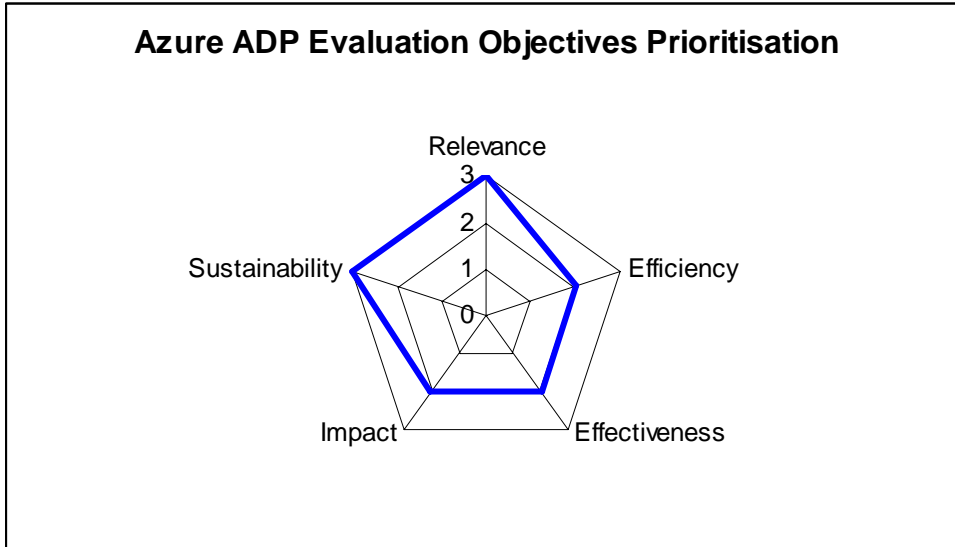
A possible evaluation conceptualization at the high level:

- Assess first whether the food security evaluation could be included as part of the ADP programme evaluation (one year early). If possible, the evaluation team needs to establish how an independent evaluation team leader could be engaged for just the food security project in Azure ADP, while the rest of the evaluation could be done in a more participatory fashion. A second option is to use an external evaluation team leader for the whole programme, but build in participatory elements (Table 2) and verify these through the evaluation TOR feedback process with Support Office “Y.” A final option is to do the programme evaluation this year with the knowledge that the food security project of ADP Azure will be evaluated in one year’s time.

- Let us assume the final option is used (do the programme evaluation this year with the knowledge that the food security project of ADP Azure will be evaluated in one year’s time):
 - **Relevance:** High score (3), given the trending toward an unstable environment and a missing programme baseline. As a result, “Relevance” questions need to be given high priority.
 - **Efficiency:** Medium (2). Efficiency measures will be more critical for the subsequent food security project evaluation, but efficiency factors need to be looked at in the general evaluation as well, particularly alternative methods of intervention delivery given the potentially changing environment.
 - **Effectiveness:** The programme has been around for a long time, so the evaluation should be able to measure changes at the outcome level of the hierarchy of objectives, but will be constrained by the lack of baseline data (though TDI s were measured). Look to assessing causality through a qualitative “contribution” approach. Need to look at how risk was managed, particularly the climate/drought issue. Assign Medium (2) rank to the Effectiveness objective.
 - **Impact:** There is a need to assess impact given the age of the programme, but the situation is constrained by the lack of baseline data. Look at the usefulness and implications of the TDI results. Consider some measures of contribution to changes at the community level through a survey (e.g. community members perceptions about what the programme has changed in the community). Assign Medium (2) rank to Impact objective.

- Sustainability: It was found that the ADP has no meaningful sustainability or transition strategies in place. Need to assess (anecdotally if not rigorously) the community in terms of progress in quality of life indicators, community capacity indicators, and transfer of responsibility indicators. One of the tasks of the evaluation could be to design these indicators and help set meaningful benchmarks for them, which could then be subsequently monitored. Assign High (3) rank to Sustainability objective.

Figure 2: Graphical Representation of ADP Azure Evaluation Objectives:



Appendix C: Evaluation Matrix Example³

Evaluation objectives	Issues	Methods	Sources
1. Assess the programme design, implementation, monitoring and evaluation processes.	1.1. Participation of stakeholders in the programme design. 1.2. Monitoring & evaluation system. 1.3 Relevance of design (does project address the needs of the target groups?) 1.4 Relevance of activities toward expected goals purposes and outputs.	1) Documentary review. 2) Field visits & observations. 3) Focus groups discussions.	1) Proposal, design, PRA Reports, semi-annual/annual narrative reports, annual operation plans, financial reports (And other grant project reports) 2) Visit & observation checklists . 3) Focus groups semi-structured questionnaire.
2. Assess the progress towards achievement of long-term (goals), medium-term (purposes) and short-term (outputs) objectives.	2.1. Achievements in terms of goals, outcomes and outputs based on logframe indicators. 2.2. Accomplishment of planned activities. 2.3 Quality of services.	1) Documentary review. 2) Field visits & observations. 3) Focus group discussions. 4) Key Informant interviews 5/ Household interviews.	1) Proposal, PRA, semi-annual/annual narrative reports, annual operation plans, financial reports. 2) Visit & observation checklists. 3) Focus groups semi-structured questionnaire. 4) Household semi-structured questionnaire. 5/ Survey questionnaire
3. Assess the potential impact in relation to relevant ministry standards.	3.1 Behavioural changes due to program implementation. 3.2 Changes in the quality of human lives. 3.3 Unexpected outcomes	1) Documentary review. 2) Field visits & observations. 3) Focus group discussions. 4) Interviews with randomly selected	1) Proposal, PRA report, semi-annual/annual narrative reports, annual operation plans, financial reports. 2) Visit & observation checklists. 3) Focus groups semi-structured questionnaire. 4) Household semi-structured questionnaire.

³ (Adapted from “Disaster Emergency Needs Assessment; Disaster Preparedness Training Program” International Federation of Red Cross and Red Crescent Societies, LEAP manual, the USAID Field Operations Manual, IFAD: A Guide for Project M&E, and the CARE M&E Guidelines)

		households.	
4. To investigate whether the resources (financial, human & materials) have been used efficiently and effectively for the well being of the target community.	4.1 Quality of structures and services. 4.2 Cost effectiveness of services	1) Documentary review. 2) Field visits & observations.	1) Proposal, PRA and other survey report, semi-annual/annual narrative reports, annual operation plans, financial reports. 2) Visit & observation checklists.
5. Assess the gender balance in planning, implementation, monitoring & evaluation as well as the access to benefits.	5.1 Accomplishment of gender equity through programme design, implementation, monitoring and evaluation.	1) Documentary review. 2) Field visits & observations. 3) Focus group discussions. 4) Interviews with randomly selected household.	1) Proposal, PRA & other survey report, semi-annual/annual narrative reports, annual operation plans, financial reports. 2) Visit & observation checklists. 3) Focus groups semi-structured questionnaire. 4) Household semi-structured questionnaire.
6. Assess the management and potentials for programme ownership, sustainability and any basis to make decision on programme transition and/or phase out.	6.1. Programme sustainability strategy. 6.2. Level of community participation.	1) Documentary review including TDI reports 2) Focus group discussions.	1) Proposal, TDI summary reports, PRA & other survey report, semi-annual/annual narrative reports, and financial reports. 2) Focus groups semi-structured questionnaire.
7. Analysis of major problems that have affected the programme (status of risks and assumptions) and analysis of the lessons learnt.	7.1 Existence and appropriateness of a risk management plan 7.2 Quality and efficiency of problems/risk management	1) Documentary review. 2) Field visits & observations. 3) Focus group discussions. 4) Interviews with randomly selected household.	1) Proposal, PRA & other survey report, semi-annual/annual narrative reports, annual operation plans, financial reports. 2) Visit & observation checklists. 3) Focus groups semi-structured questionnaire. 4) Household semi-structured questionnaire.

Appendix D: Types of Data Collection

Note: this list is not exhaustive.

- **Document Review (Secondary Data)**

This method can serve many purposes, including the gathering of baseline information. Make sure you are clear about the questions you wish to answer and what type of information you need for this. List all possible sources of existing information (project documentation, government records, organization reports, university studies, etc). Prioritise those that are most likely to provide useful information in a cost and time-efficient manner. The idea is not to read everything, especially if there is a glut of information, focus on main points only. Collect the documentation and check its reliability: note contradictory evidence and analyse it in terms of the questions you are trying to answer. Identify which information gaps you still have or where contradictory evidence needs to be clarified. Select another collection method, such as questionnaires or interviews, to fill that gap.

- **Onsite Visual Inspection**

For an experienced observer, this is an excellent way to become familiar with a disaster situation. This method does not stand alone, however, and other assessment methodology should be used in conjunction.

On-Site Visual Inspection Tasks:

- Observe people's physical condition and activities; ask questions.
- Visit homes or shelters, water sources, clinics, distribution centres.
- Observe children, the elderly, and the sick.
- Observe the daily lives of women (use women as interviewers).
- Observe the services, vehicles, and sanitation systems.
- Make sketches, take photographs, or use videos. Photos, video footage, and even hand sketches are extremely useful in communicating to others the reality of the situation.

- **Interviews**

This technique can include interviews with individuals or groups. Loosely structured interviews are conducted with key informants in the government, NGOs, and within particular groups of the affected population including: local officials, local (non-official) leaders and experts, local media sources, and leaders of the affected population. The focus is on obtaining factual information that is crosschecked with other sources. Mix in open questions as well as yes/no and multiple choice.

- **Focus Group Discussions** with members of the affected population, local leaders, or officials can be useful for crosschecking information with a number of respondents. They are often valuable for getting a sense of the issues affecting all members of a group (rather than individual issues). Over-reliance on group interviews, however, may leave information gaps because there may be issues that are not discussed in groups or because some people may not be represented in public groups (like women).

Groups should be no bigger than 4-8 people, and depending on the purpose of the discussion, can be comprised of either a heterogeneous group (e.g., a mix of men, women, ethnic groups, etc.) or a homogeneous group (e.g., all women). The timeframe for discussion should be determined beforehand, and the facilitator should have minimal engagement in the discussion. A note-taker should be present to record the results.

- **Questionnaires and Surveys**

This method allows a surveyor to gain information from a large number of people in a structured way according to specific questions, often in ways that allow for statistical analysis. These can range from being very simple to quite complex, which makes them useful in both relief and rehabilitation/development contexts. They can follow a very specific and structured set of closed questions (yes/no or multiple choice questions) or they can also include open-ended questions, such

as semi-structured interviews. Fixed-choice or fixed-response questionnaires are good for gathering data that needs to be analysed statistically, while open-ended or free-response questions can be particularly good for determining people’s feelings and attitudes.

A Note on Translation

The reliability of data collected by qualitative means such as focus group discussions and questionnaires and surveys relies heavily on good translation. If the translators do not have a clear and strong understanding of the material and the *purpose* of the questions, then the exercise will not yield meaningful results.

Examples of Closed Versus Open Questions

CLOSED QUESTIONS	OPEN AND FOCUSED QUESTIONS
Have you been registered for disaster relief for your family? Are you currently living in a camp or are you able to still live in your home after the disaster?	What relief supplies do you need? Which have you already received? What kind of shelter assistance do you and your family require?
Do you grow enough food to satisfy your family’s needs?	What staple food crops do you grow? Do you have enough food to feed your family today/this week?
How often do you attend the maternal child health clinic with your children?	Do you have children under five years old? How often is the clinic conducted in your village?

Example of a Mini-Survey

	Home undamaged by disaster	Home partially damaged by disaster	Home completely destroyed by disaster
Household 1	X		
HH 2			X
HH 3	X		
HH 4		X	

Questions must be carefully formulated and the survey team should agree on who should be questioned and how many people should be included in the sample before hand. Ensure that a fair sample of the population is questioned, including men, women, children, and vulnerable groups.

- **Sampling**

Sampling is a method by which a generalization about an entire population is made based on the characteristics of a subset (or sample) of the population. Attributes or proportions of a population are estimated through interviews or surveys with a representative section or sample of the population. Information collected through sampling methods includes written questionnaires and interviews.

There are two types of sampling techniques: probability and non-probability. Probability sampling is based on rigorous statistical methods. It is expensive and time-consuming to implement and requires special training to be used correctly. Non-probability methods are commonly used and rely on interviews with those who are most accessible (convenience sampling) or with individuals that are believed to be representative of the population of interest (purposive sampling).

- **Case Study**

This method is meant to give a snapshot of a particular family or individual’s situation over a period of time and can be used for assessment, baseline, and monitoring purposes. It documents the life story or sequence of events over time related to a person, location, household or organization in order to obtain insight into a project’s effect and to learn about people’s experiences, dreams, and obstacles for future planning. Case studies can add life to hard data and allow for in-depth understanding of the context and human factors behind summarized data collected through other means. The team should agree beforehand as to how the subject will be selected. Questionnaires, semi-structured interview questions, and other relevant methodologies will need to be employed.

- **SWOT Analysis (Strengths, Weaknesses, Opportunities, and Threats)**

This method could be used to determine strengths, weaknesses, opportunities, and threats in relation to a project or group and how such a situation could change over time. This method is useful when qualitatively assessing, for example, the services provided by the project, relationships between project stakeholders and the organizations of the implementing partners, local groups, and the project team itself.

The group must define, discuss, and record as many factors as possible for each heading. Emphasize that strengths and weaknesses refer to internal aspects of the group, while opportunities and threats can be looked at in terms of internal or external factors affecting them.

Strengths—Those things that are working well in a project or situation. The aspects people are proud to talk about

Weaknesses—Those things that have not worked so well.

Opportunities—Ideas on how to overcome weaknesses and build on strengths.

Threats—The things that constrain or threaten the range of opportunities for change.

- **Rapid Social Organization Profile**

Information about existing social groups, including LNGOs, in an affected community can help the WV response team to quickly identify potential implementation partners and key stakeholders in the response.

Group Name	Size	Gender of members	Age	Admission rules	Activities	Date group started	Location	Key contact	Sectoral expertise

- **Detailed Critical Sector Assessments by Specialists**

This involves technical inspections and assessments by experts. It is required in sectors such as health and nutrition, water and sanitation, shelter and infrastructure. Critical sector assessments may be compiled from reports by specialist teams from outside World Vision or from primary data gathered by WV sector specialists themselves.

Appendix E: Team Members and Roles example table

Evaluation Phase	Role	Primary Task
Planning	Program Manager	<ul style="list-style-type: none"> • Coordinate collection of information for terms of reference • Recruit Team Leader • Recruit other team members in consultation with the Team Leader • Advise the Team Leader on evaluation priorities from a program perspective • Introduce Team Leader to evaluation partners
	M&E Specialist	<ul style="list-style-type: none"> • Advise manager in writing the terms of reference
	Team Leader	<ul style="list-style-type: none"> • Get input on purpose and objectives from partners • Prepare detailed plan for data collection and analysis to achieve evaluation objectives
Data Collection and Analysis	Team Leader	<ul style="list-style-type: none"> • Supervise data collection and analysis
	Technical experts as needed	<ul style="list-style-type: none"> • Ensure that the quality of data is acceptable • Supervise analysis and interpretation of technical data
	Translators as needed	
	Focus group leaders and recorders as needed	<ul style="list-style-type: none"> • Lead focus groups • Record responses and qualifying observations • Prepare summaries of responses
	Program staff and community representatives	<ul style="list-style-type: none"> • Provide input on what the findings mean to them • Develop practical recommendations
Reporting and Follow up	Team Leader	<ul style="list-style-type: none"> • Draft the report and circulate it for feedback • Incorporate feedback into the report
	Team members	<ul style="list-style-type: none"> • Provide critical feedback on draft
	Translators	<ul style="list-style-type: none"> • Translate portions of draft as needed to get input from people that speak the local language
	Program staff	<ul style="list-style-type: none"> • Provide feedback on the understandability of the report
	Evaluation partners	<ul style="list-style-type: none"> • Provide feedback on how well their objectives have been met