

OFFICE OF THE INSPECTOR GENERAL

IOM EVALUATION GUIDELINES

January 2006



IOM International Organization for Migration

TABLE OF CONTENTS

	Page
Foreword	1
1. Introduction: Building an Evaluation Culture in IOM	2
2. Understanding what Evaluation is	4
2.1 Evaluation Types	4
2.2 Evaluation Methodologies	7
2.3 Other Related Notions	9
3. Managing and Conducting Evaluation: Norms, Standards and Criteria	11
3.1 Norms and Standards for Evaluation	11
3.2 Evaluation Criteria	14
3.3 The Reference Data	16
4. Utilizing, Planning and Preparing Evaluations	18
4.1 Utilizing Evaluation: Accountability, Learning and Promotion	18
4.2 Planning an Evaluation	21
4.3 Preparing and Organizing an Evaluation	22
Annex 1 Bibliographical References	25
Annex 2 Other Types of Evaluation	28
Annex 3 Some Questions related to Evaluation Criteria	30
Annex 4 Evaluation Report Format	34
Annex 5 Self-Evaluation	36
Annex 6 Terms of Reference Format	46

FOREWORD

In recent years IOM has demonstrated its capacity to meet numerous worldwide demands and adapt to new challenges in the evolving field of migration management. The dramatic increase in volume and scope of our projects and programmes bears witness to this. But quantity is not the only valid measure of what we do. Quality and cost effectiveness are also key benchmarks. Evaluation plays a key role in ensuring accountability to donors by demonstrating that work has been carried out as agreed and in compliance with established standards. An evaluation culture must be fostered in order to gain the full benefit from what can be learned from what we do and how we do it, and ensure that we continue to build on our recognized strengths of flexibility, reliability and creativity.

IOM's Member States and Donors support strong oversight functions capable of reviewing activities and providing performance feedback. I created the Office of the Inspector General, which encompasses the Evaluation function, to meet this demand and, in order to ensure transparency IOM Evaluation's work is made available through its webpage on IOM's website. Just as IOM continues to increase and diversify its activities, so IOM Evaluation needs to adapt to the changes. OIG has produced new evaluation guidelines to replace those of 1998, an initiative that I fully support.

The decision to conduct an evaluation is left to the responsibility and judgment of the project and programme managers. The guidelines present the benefits clearly: evaluation not only covers learning and accountability, but can be used to promote IOM's work, reinforce partnerships and bring innovation to its activities.

The primary users of the guidelines will be the Heads of MRFs, Chiefs of Missions and Project and Programme Managers but I strongly encourage you all to familiarize yourselves with the guidelines to know more about evaluation in IOM and enable you to discuss it knowledgeably with partners and donors.

Brunson McKinley
Director General

“The damage does not come from the finding of an evaluation, but from people trying to hide it.” (Michael Q. Patton)

1. INTRODUCTION : BUILDING AN EVALUATION CULTURE IN IOM

Efforts to promote evaluation in IOM already started at the beginning of the 1990s. The first IOM evaluation guidelines were published in 1992, briefly presenting evaluation and proposing an evaluation report format, but not including a clear evaluation policy and strategy. The IOM's traditional movement services and resettlement programmes, provided under contractual basis and accounting for more than 70% of its operational budget at that time, were regularly reviewed through internal and donors' audits. Evaluations were conducted in response to specific needs and requests.

In 1998, IOM adopted an evaluation strategy to more actively promote the use of evaluation in its programmes and projects. The scope of activities carried out by the Organization was rapidly expanding to include a broad diversity of projects covering various aspects of migration management. In 1999, the Organization published its second 'Evaluation Guidelines' containing also information on how to perform self-evaluation.

The Office of the Inspector General – OIG, created in 2000, was tasked to continue the promotion of evaluation in IOM and to review the evaluation policy, strategy and guidelines when deemed necessary. In that perspective, OIG has been examining the relevance and effectiveness of the 1999 guidelines and considered, in 2004, that it would be opportune to update them to include the new evaluation trends agreed upon and implemented by the international community, and to adapt them to IOM's continuously evolving work.

OIG also decided to revise specific evaluation requirements, in particular the use of self-evaluation and the systematic inclusion of an evaluation exercise in IOM project/programme documents. OIG noted, for instance, that several projects in IOM were not too complex and were implemented for a one year period with relatively low budgets, which put into question the validity of a costly evaluation and made it inappropriate to plan a full fledged evaluation if the project was not extended. Detailed donor's reports were often deemed sufficient to draw conclusions on the performance and achievements of IOM smaller projects.

It should also be noted that OIG introduced the internal rapid assessments into its core tasks that examine the performance and achievements of projects from a neutral outsider's perspective. Even if kept internal, the reports bring useful information and lessons on the implementation of projects, which can effectively replace small evaluations or self-evaluations in several cases.

OIG also concluded that some sections of the 1999 guidelines had not been well developed and that it was necessary to define and explain the evaluation concepts in more detail in order to avoid confusion in the terminology used by IOM field offices and in the definition of key terms. Other sections were, on the contrary, too detailed and not properly adapted to the targeted users, for instance a full section covering the conduct of the evaluation that is, in most cases, the work of consultants.

In order to promote evaluation and conduct self-evaluations, IOM had in the past organized regular evaluation training sessions during the Project Development training courses targeting project developers and Chiefs of missions. In 2000, following a review of the IOM overall training strategy, the Project Development training was restructured and the timeframe shortened, and the evaluation module was no longer a main tool for the promotion of evaluation. New evaluation guidelines are an opportunity to fill the gap of training IOM staff and meeting the expectations of IOM staff tasked to discuss and use evaluation.

The new guidelines will therefore concentrate on increasing IOM staff capacity to think in 'evaluative' terms for a promotion of an evaluation culture, and on providing technical references that can clarify what evaluation is in order to facilitate discussions with partners and the donor community when developing projects or discussing the implementation of evaluation exercises. As a basic principle, OIG considers that constraining and mandatory instructions for the use of evaluation do not help towards changing its negative perception as a 'policing' tool, and will try instead to have it perceived as learning and rewarding exercise. It is hoped that the positive aspects of evaluation will be recognized.

Section 2 of the guidelines will propose several definitions related to the concept of evaluation. The section will also include a description of various types and methodologies, which will help to frame the complexity of the notion. Section 3 will present the evaluation principles, norms and standards, as well as the various criteria and concepts used in the framework of the evaluations of development assistance and humanitarian actions. Section 4 is more practical and will try to explain why evaluation could be useful and how to use it effectively. It will also provide guidance on how to include evaluation in the project development and how to prepare it.

“But this era of endless debates on evaluation utilization should now end. New realities ask for, indeed demand, a different conceptualization about evaluation utilization. We are in a new period where ever-accelerating political and organizational demands are reframing our thinking about the definition of what, fundamentally, constitutes evaluation and what we understand as its application.” (Ray C. Rist)

2. UNDERSTANDING WHAT EVALUATION IS

‘Evaluation’ is used in various contexts without precise delimitations and definition; its meaning seems evident. A broad definition often applied is: *the process of forming an idea of the amount, number, value or merit of something*. However, the need for more precise definitions and knowledge of the concept emerged as the theory of evaluation evolved and became more and more complex. In the field of development assistance and humanitarian action, the most common definition¹ is:

‘the systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfillment of objectives, development efficiency, effectiveness, impact and sustainability. [...] Evaluation also refers to the process of determining the worth or significance of an activity, policy or programme’.

Diversification of methodologies for conducting evaluation exercises in the last fifty years also led to the elaboration of more specific definitions, sub-categories or types of evaluations. An extensive literature exists on the topic covering all kinds of situations in which evaluation is used. It is important to take into account that evaluation used for IOM programmes or projects is only a small portion of this big ‘territory’ and is less complex technically and in scope than evaluations implemented, for instance, at national levels in the framework of national education programmes or national health policy.

2.1 Evaluation Types

The types of evaluation can be grouped under four categories based on the objectives and recipients of the evaluation, the timing of the evaluation, the ‘evaluator’ and the technicality or methodology used. Any evaluation could also be a ‘mixture’ of the different categories.

- 1) A first distinction can be made between two approaches of evaluation related to the timing and recipients of the evaluation: *formative* and *summative* evaluations. A **formative evaluation** may be defined as an evaluation conducted during implementation for improving performance. It is intended for managers and direct actors to help them to redress and improve the project or programme according to participants’ suggestions and needs. A **summative evaluation** is conducted at the end of a project or programme and for actors not directly involved in the management of the implementation (for instance donors). It provides insights about

¹ OECD, Development Assistance Committee – DAC, Working Party on Aid Evaluation, *Glossary of Key Terms in Evaluation and Results Based Management*, 2002.

the effectiveness of the project or programme and gives then the opportunity to use best practices identified during the evaluation.² IOM rarely makes the distinction when discussing evaluation; both types are used and sometimes the evaluation product ends up being a mixture of both (for instance conducted during the implementation but for actors not directly involved).

- 2) Another distinction made is based on the timing of the evaluation exercise and four main types can be mentioned: *ex-ante*, *mid-term*, *final (or terminal)* and *ex-post*. An **ex-ante evaluation** is performed before implementation of assistance. However, it is not often used as *appraisals*, *needs assessments* or *feasibility studies* are more practical and common tools. A **mid-term evaluation** is carried out approximately at the middle of the implementation of the project or programme. It is usually for projects or programmes that are planned for lengthy periods and is used to assess adjustments that can be brought to enable the project or programme to achieve its goals. A **final or terminal evaluation** is carried out at the end, or close to the end, of the project or programme when all aspects can be considered. This is the most common type used in IOM. The fourth option is the **ex-post evaluation**, which is done some time after the project's or programme's end. It is envisaged if there is a need to focus on the impact and sustainability, but could be problematic as often the projects or programmes are already closed. Beneficiaries and staff having participated in implementation are no longer easily reachable.
- 3) A third distinction is made according to the person(s) who conduct(s) the evaluation exercise, with two main categories: *internal* and *external*. The notions of *independent*, *self-evaluation* and *joint evaluation* are also discussed here. An **internal** evaluation is conducted by a unit and/or individuals belonging to the organization, government, department responsible for implementing the project or programme. It is an **independent internal** evaluation if conducted by somebody who did not directly participate to the implementation. It is an **internal self-evaluation** if done by those who are entrusted with the design and delivery of the project or programme. In IOM for instance, OIG is conducting independent internal evaluations. An **external** evaluation is conducted by someone outside the donor and implementing organizations. External evaluations often require the recruitment of consultants and therefore are certainly the most expensive formula. They are independent evaluations. **Joint evaluations** can be a mix of internal and external categories; their specificity lies in the fact that both donor agencies and implementing partners participate. As mentioned in the OECD/DAC glossary, "there are various degrees of 'jointness' depending on the extent to which individual partners cooperate in the evaluation process, merge their evaluation resources and combine their evaluation reporting". It is evident that the organization of joint evaluations is more complex than a single external or internal evaluation due to the coordination aspects between interested partners. There are examples of joint evaluations in IOM.
- 4) The fourth group is identifying the types of evaluation according to technical specificities and scope. This is the most diversified and these guidelines will not list the definitions of all existing types. The most common ones will be presented in the core text and others will be added in Annex 2. A brief listing of the terms highlights that diversity: *Programme evaluation*, *Project evaluation*, *Thematic evaluation*,

² Most of the definitions mentioned in Section 2.1 are from, or adapted from OECD/DAC glossary papers, including October 2000, November 2000, April 2001 and 2002 versions. See also Annex 1 for further references.

Impact evaluation, Participatory evaluation, Process evaluation, Country-Programme evaluation or Country Assistance evaluation, Result-based evaluation, Strategic evaluation, System-wide evaluation, Partial system evaluation, Single agency response evaluation, Single agency/single project evaluation, Real-time evaluation, Sector evaluation, Cluster evaluation, Empowerment evaluation, Stakeholder evaluation, Quasi-experimental impact evaluation, Meta-evaluation, Cross-section evaluation, Incorporated/built-in evaluation, In-depth evaluation, Theory based evaluation, Democratic evaluation etc.

The 'standard' types used in IOM are the **Programme and Project evaluations** that examine respectively a set of activities brought together to attain specific global, regional, country or sector assistance objectives, and an individual activity designed to achieve specific objectives within a given budget and time period. **Thematic evaluations** are also conducted in IOM, mainly by OIG, and examine selected aspects or crosscutting issues in different types of assistance (poverty, environment, gender, etc.).

The two following types are not often used in IOM but are mentioned here as they may lead to confusion. Speaking of the 'evaluation of the impact' is very common in the discussions about the performance and achievements of a project or programme, and impact is part of the criteria that are used in evaluation (see Section 3). The analysis of the impact is often deducted from various observations and findings without necessarily measuring it. However, an **Impact evaluation** is different and requires specific methodologies and precise and systematic technical steps in order to elaborate valid and verified conclusions and recommendation. It may be defined as a systematic identification of the effects – positive or negative, intended or not – on individuals, households, institutions, and the environment caused by a given activity. It helps to measure outcomes and impacts of an activity on a target population and distinguishes these from the influence of other external factors. The 'causality' linkage in the impact evaluation is also a complex analysis. The budget for conducting an impact evaluation can also be high, requiring detailed surveys on broad population samples, and the exercise time-consuming. It may be inappropriate if there is a need for rapid results. A basic principle to apply before choosing an impact evaluation is that the benefits of the evaluation should outweigh their costs and limitations.

Confusion also exists between 'participatory' techniques for conducting an evaluation, or writing a project document, and **Participatory evaluation** which may be defined as an evaluation method in which representatives of agencies and stakeholders (including beneficiaries) work together in designing, carrying out and interpreting an evaluation. The collaborative effort deserves to be underlined but it also brings organizational constraints that render the exercise relatively complex. Participatory techniques consist for instance of focus group discussions or preparatory meetings.

The following evaluation types are relatively common in international cooperation activities and deserve to be mentioned. A **Process evaluation** examines the internal dynamics of implementing organizations, their policy instruments, their service delivery mechanisms, their management practices, and the linkages among these. A **Country-programme** or **Country-assistance evaluation** is more common in UN agencies and bi-lateral assistance that use country programming approaches, and is defined as an evaluation of one or more donor's or agency's portfolio of development

actions, and the assistance strategy behind them. Finally, **Result-based evaluations** are related to the Result Based Management (RBM) methodology, or the 'managing for results' approach, and look at the performance and achievements of development assistance through results oriented approach and performance indicators.

There are still many discussions on how to make the best use of evaluation in the RBM theory, or even if evaluation is necessary when using a RBM model. The RBM approach is used for complex programmes implemented at the country level and that require a set of indicators to properly monitor and evaluate the assistance. In IOM, the projects and programmes are less complex and the approach, when applied, ends up being very similar to the logical framework methodology, *logframe approach*, used in IOM.

2.2 Evaluation Methodologies

The methodologies used in evaluation are diversified and have been subject of extensive literature. Not all are specific to evaluation and some are also used for research or monitoring.

Two major methodological approaches can be first identified: the *quantitative* and the *qualitative* analysis. The **quantitative analysis** relies mainly on statistical data and their basic interpretation. The main challenges of the quantitative analysis are certainly how to 'humanize' the numbers, how to consolidate data and their meaning, and how to make the report concise and interesting. The **qualitative analysis** refers more on 'judgemental' mechanisms based on observational data. One of the main challenges is the validity and credibility of the observation and judgement made. In that perspective, basic principles and standards for conducting evaluations are important elements in rendering the analysis valid and relevant. Aside from the specialized literature that clearly delimitates the approaches, the common practice in development and humanitarian assistance evaluation deals most of the time with a combination of both methods.

For both quantitative and qualitative analysis, a large set of methodologies can be recorded. The most common methodologies used for IOM evaluations, as well as in many UN agencies, are limited to the following: *Documentation review*, *Direct observation*, *Interviews*, *Focus groups*, *Questionnaires and surveys* and *Case study*, each of them having sub-categories and specific procedures.

The **documentation review** provides useful baseline information and an historical perspective of the project or programme. It includes written documentation, but also videos, electronic data or photos. The documentation review is not limited to the project's or programme's activities and can cover any information on the strategy or global approach, closely related or not.

The **direct observation** consists mainly of what can be seen during field travels and project sites visits. It can be subjective as human vision and perception is selective, based on fields of interest or cultural sensitivities. It is important for the evaluator to use it knowledgeably and cross-check facts and elements collected.

Interviews are also an important methodology for collecting data and information. There are various interviews' techniques. *Structured interviews* or *strict interviews* require an *interview's guide*, following strictly and systematically questions prepared in advance. If one advantage is that the same questions are asked of all interviewees, implying a diminished risk of 'manipulating' the questions, they are not exploring all the possibilities for new questions that can emerge during the interview as is the case with the *semi-structured interviews*. A third option is called the *conversational interviews*, which also follows a line, but remains very open in asking questions or in formulating them. Interviewing a group of persons can be done through *brainstorming* or *group discussions*. At times, without making a differentiation, they are called *Focus groups*.

Focus groups, however, require specific techniques and the selection of the groups should follow precise criteria and procedures. Two persons should manage the discussion, one taking notes and the other guiding the debate and looking at people's behaviours. Focus groups may also be used during the preparatory phase of research, especially to determine the content of a survey and the appropriate sampling methods.

Questionnaires and surveys are important tools to gather data from a large number of people in a structured way that often allows for statistical analysis. They can be short and simple, with 'yes-no' answers, or very long and complex using open questions, numerical scales or 'agree-disagree' scales. The more extended the survey is, the more complex its analysis and interpretation are. It is important to properly prepare the questionnaire/survey and formulate the questions based on the objective(s) and scope of the survey. The analysis will be greatly facilitated. *Before and after surveys* are often used in IOM especially to measure the awareness raising of IOM information campaigns and counter-trafficking activities.³

A questionnaire or survey can be sent to the overall target population if the population is small, but sometimes this is not possible due to its size or because it is too disseminated among different countries (especially in the field of migration). Sampling methods play an important role to have a representative sample, statistically validated for interpreting the findings of the survey. There are two major categories of sampling methods: *random sampling* and *qualitative sampling or purposeful sampling*.⁴ The first one is used when there is a need for generalizing from the sample, and the second one when the sample is established for a specific purpose, not necessarily for generalizing.

In IOM, as in most evaluations conducted by the UN or other international cooperation/humanitarian agencies, qualitative sampling is used due to the specific objectives of a programme or project. A preliminary analysis of the sample is made in order to examine the restrictions in formulating conclusions. Then the information is cross-checked with other data, for instance those recorded by other existing surveys, in order to draw relatively valid conclusions. If a correct sample is important for the validity of the survey, sometimes it is not possible to strictly follow the methodological

³ See also: IOM, *Research and Evaluation methodology for Mass Information Activities*.

⁴ Other techniques include for instance *stratified samples* where a break down by age, gender, social origin etc. is made, or *systematic sampling* where a sample is selected at predetermined intervals. *Cluster sampling*, *quotas sampling*, *homogenous sampling*, *maximum variation sampling* are other examples. Close to 20 sampling strategies can be recorded.

requirements, especially with large target populations. The main reasons are the costs for such surveys.

The last methodology explained here is the **case study**. It is the selection of a life story, specific events or of projects/programmes activities that can give an insight of a project's effectiveness and impact. The case study in isolation does not prove anything, it helps to illustrate a data and find commonalities. Only when adding cases, then one can extrapolate some principles.

As already mentioned, there are several other methodologies that can be used for conducting complex evaluations, some adapted to very specific situations, and that are not necessarily used in IOM evaluation work. Just to highlight the variety, here are other examples: *Triangulation, Outcome mapping, If-then, Performance logic chain assessment, Participatory process, Social mapping, Biophysical measurement, SWOT (Strengths, Weaknesses, Opportunities and Threats), Dreams realized, Impact flow diagram etc.*⁵

As underlined in the beginning of the section, many, and perhaps most evaluations actually use mixed approaches. **Mixed methods** help compensate the technical gaps that could be encountered in using only one methodology as explained, for instance, in the discussion above on quantitative versus qualitative methods.

2.3 Other Related Notions

Some notions are on occasion used instead of evaluation or are closely related to evaluation but need to be differentiated mainly for methodological reasons. The definitions below will help to avoid confusion in the terminology used.

Evaluation has a **research** component as it is also a sum of work which tends to the discovery of new knowledge or to developing existing knowledge. The methodologies and technical tools used for research are often similar to the ones used in evaluation (quantitative and qualitative analysis, survey techniques, sampling and inquiry methods etc.). There are also various categories of research based on different approaches (*basic research, applied research, action research etc.*). Research is a learning process but without the control and accountability components that are specific to evaluation. Another difference is that research very often does not include 'stakeholders', contrary to evaluation.

Another close 'friend' of evaluation is **audit** that may be defined as an activity of supervision verifying whether the existing policies, norms and instruments are being applied and used adequately. Audit also examines the adequacy of organizational structures and systems and performs risk assessments. The audit focuses on the accountability and control of the efficient use of resources, without the learning or research dimension that can be found in evaluation. Some auditors will however disagree on the fact that there is no 'learning' in audit, especially when they refer to the **performance audit**. If both the evaluation and performance audit share the same objective, to improve the quality and implementation of a programme or project, the

⁵ For more information and definitions on methodologies, see in particular: IFAD, *A Guide for Project M&E*, Annex D, 2002. For further technical bibliographical references, see Annex 1.

performance audit focuses on the effectiveness and does not include the analysis of the relevance, sustainability and impact of the activities.⁶

Another term that is often used interchangeably with evaluation is **review**. According to the OECD/DAC glossary, a review is “an assessment of the performance of an intervention, periodically or on an ad hoc basis. A review is more extensive than *monitoring* but less than evaluation”. An evaluation is more comprehensive and technically more constraining than a review.

Monitoring can be defined as a continuous function that uses systematic collection of data on specified indicators to provide management and stakeholders with indications on progress and achievement of objectives, and to take decision to improve performance. Sometimes reference is made to **on-going evaluation**.

Needs assessments and **appraisals** are tools enabling decision makers to choose and decide between optional activities, and to refine the final design of a project or programme. They are assimilated to ex-ante evaluations and take place before implementation. **Rapid assessments** or **rapid appraisals** are conducted during implementation and are quick ways to gather views and feedback from beneficiaries and stakeholders on an on-going activity, project or programme in order to respond to decision-makers needs for information. They are often less expensive but also less rigorous than formal evaluations or surveys. There is no specific methodology and techniques linked to them; the focus is on ‘rapid’. In IOM, rapid assessments focus mainly on what the organization promises to deliver within a given timeframe and budget, and how well IOM is actually delivering. The Oversight Officer in OIG conducts rapid assessments of IOM projects and programmes and the reports are internal.

Very similar to rapid assessments in terms of methodological approach and objectives, **Real time evaluations** are conducted in emergency situations. An emergency requires immediate large-scale response, flexibility and adaptation to rapidly evolving situations, often without sufficient baseline data to draw comparisons. The duration is short and the workload created by the evaluation on the staff working in the emergency is minimal. Real time evaluations are also expected to produce an immediate feedback and contribution.

OIG also includes two other functions: *Inspection and Investigation*. **Inspection** can be defined as a general examination seeking to resolve a particular problem, or to identify vulnerable areas and malfunctions and to propose corrective actions. An **investigation** is a specific inquiry or examination of a claim of wrongdoing and provision of evidence for eventual prosecution or disciplinary measures.

⁶ For the definitions of effectiveness, relevance, sustainability and impact, see Section 3.

“We need to establish credibility (i.e., of both expertise and objectivity) if evaluation findings are to be used either for enlightenment or in policy making. [...] We need to let evaluation be evaluation – that is to keep its skepticism about the conventional wisdom, its meticulousness about measuring achievements, its willingness to be persistent about getting the information out, and its dedication to democratic reform on the basis of knowledge.” (Eleanor Chelimsky)

3. MANAGING AND CONDUCTING EVALUATION: NORMS, STANDARDS AND CRITERIA

The theories on evaluation developed basic principles, norms and standards in order to guarantee quality, credibility, impartiality, reliability, objectivity, validity, and transparency. These concepts have been and are still subject to numerous debates and are difficult to contain into a rigid and delimited frame due to their complexity and subjective dimension.

Evaluation analyses the relevance, performance and achievements of a project or programmes through a set of well defined criteria. Notions such as ‘indicators’, ‘baseline data’ and ‘benchmark’ are also presented under the section.

The information under the section may also help to assess and guarantee the quality of an evaluation and to understand what to examine when conducting an evaluation.

3.1 Norms and Standards for Evaluation

There is no clear delimitation of what precisely encompass ‘norms’ and ‘standards’ for evaluation and sometimes they are mixed; reference is also made to ‘principles’ for evaluation, ‘guiding principles’, ‘best practice guidelines for evaluation’, or ‘logic and values of evaluation’. Standards and criteria common to the research field may also be applied to evaluation work, and sometimes used as reference. This section will however not go into too much detail as it is not the main purpose of the guidelines. It will offer some basic references that help to better frame the complexity of managing and conducting evaluation in order to guarantee quality, credibility, impartiality and usefulness.

The American Evaluation Association⁷ identifies four groups of programme evaluation standards listing in total 32 standards. The **utility** standards are intended “to ensure that an evaluation will serve the practical information needs of intended users”. They include among other requirements a proper identification of the stakeholder, the evaluator’s expertise through an appropriate selection, a relevant selection of information to be collected, clarity in the report with clear description of the perspectives, and selection of procedures and rationale to interpret findings.

The **feasibility** standards are intended “to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal”. They include among others the need to select

⁷ See AEA Website for further reference: <http://www.eval.org/Publications/publications.html>.

practical methodologies and procedures, with minimum disruption, a political viability with proper plans to obtain the cooperation of various interest groups, and cost-effectiveness where the benefit of evaluation should outweigh the costs.

The **propriety** standards are intended “to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results”. They include among others the need for formal agreements among parties on who does what, how and when, the respect of human dignity and the rights of human subjects so that participants are not threatened or harmed, complete and fair assessment, and proper disclosure of findings taking into account pertinent limitations.

The last group is the **accuracy** standards that are intended “to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the programme being evaluated”. They include among others a proper description and analysis of the context, a detailed presentation of the objectives, scope and procedures of the evaluation, defensible information sources, valid, reliable, systematic information, justified conclusions and impartial reporting.

The American Evaluation Association also established the ‘Guiding Principles for Evaluators’ that are directed towards the choice of the person(s) who should perform the evaluation and how he/she should conduct it, and more generally behave. Some of them are closely related to the standards described above; they include systematic inquiry, competence, integrity/honesty, respect for people, and responsibilities for general and public welfare.

The United Nations Evaluation Group (UNEG), of which IOM is a Member, developed recently two papers presenting respectively the ‘*Norms for Evaluation in the UN System*’ and the ‘*Standards for Evaluation in the UN System*’.⁸ In addition to the requirements regarding the evaluators and the evaluation exercise itself, which are in line with what is already described above, they include norms and standards for the management of the function of evaluation at the level of the organization, including the institutional framework, the policy, and the reporting to the heads of the organization and governing bodies. In the case of IOM, that category of norms and standards applies more specifically to the OIG/Evaluation’s role.

Two norms presented in the UNEG paper are often discussed when preparing, conducting or using evaluation, and deserve to be mentioned: *Impartiality* and *Transparency*. **Impartiality** is defined as “the absence of bias in due process, methodological rigour, consideration and presentation of achievements and challenges. It also implies that the views of all stakeholders are taken into account. [...] Impartiality provides legitimacy to evaluation and reduces the potential for conflict of interest”. Impartiality is often discussed together with the notion of *independence*.⁹ If they are closely related, the possibility exists to have, however, a non-independent evaluation (for instance a self-evaluation) which would be fully impartial. Other notions such as *neutrality* or *objectivity* are also used to describe what can be recorded as elements of the intellectual rigour and professional integrity required for an evaluation of quality.

⁸ Both documents are available under the IOM Evaluation Web page for more detailed references.

⁹ See Section 2.1 on independent evaluation.

The definition of **transparency** is restricted to the spheres of *consultation* and *public availability*, in the sense that evaluation requires consultation with the major stakeholders during the entire process, including for the preparation of the terms of reference, and the reports to be made public. In some literature, the concept is also applied to the process and methods for conducting the evaluation that should be transparent. Transparency when applied to the content of the report raises, however, more questions and sometimes 'authorizes' restrictions for practical and effective reasons.

The underlying question is: can the evaluator, in case of poor results, write anything against a government, an organisation or a group of individuals/stakeholders in the name of transparency, even if the statements are well supported and perhaps justified? The answer is 'no' for the simple reason that the development and humanitarian assistance evaluation is part of a 'political sphere', or 'political game', and cannot work in complete isolation. Notions such as *diplomacy* or *political wisdom* play an important role. If not properly balanced, the end result could be that the evaluation will not be used at all and even denigrated, as being too confrontational. This is not the goal of an evaluation.

Evaluation requires the capacity to 'negotiate' different realities, different reading of the world, with various and sometimes diverging cultural sensitivities, without offending. Full acceptance of the content by all parties, also if self critical, is part of the quality of an evaluation; the evaluation should be an asset for those being evaluated. The feasibility standards described above and the notion of *usefulness* are of importance here.

The UNEG paper does not include *usefulness* as a norm for evaluation, but is however implicit. The OECD/DAC considers **usefulness** as a principle for guaranteeing quality: "to have an impact on decision-making, evaluation findings must be perceived as relevant and useful and be presented in a clear and concise way. [...] The evaluation process itself promotes a further clarification of objectives, improves communication, increases learning, and lays the groundwork for follow-up actions".¹⁰ The utility standards are very similar to the principle of usefulness, but more restricted to the methodological process.

A last common principle that deserves to be mentioned in the guidelines is **credibility**. Credibility is also part of the elements of intellectual and ethical rigour and professional integrity required for an evaluation of quality. All the standards and norms described above contribute to the credibility of the evaluation, including among others the expertise and impartiality of the evaluator, the mutual trust between the evaluator and those evaluated, the degree of transparency, the clarity of the report with a description of both successes and failure, etc. The lack of credibility undermines the use of the evaluation.

IOM is recognized for its flexible but also innovative approach to migration management. Designing new programmes, establishing new strategies for responding to the strongly evolving realities in the field of migration, requires innovation and

¹⁰ See OECD/DAC, *Principles for Evaluation of Development Assistance*, 1991.

creativity. **Creativity** is not recognized as a principle, norm or standard in the evaluation literature, but some authors are arguing that creativity could be an asset.¹¹ Evaluation being an exercise focussing not only on accountability but also on learning can generate new knowledge too. Creativity in evaluation can lead to innovative working hypotheses, open new fields of interest, new parameters of action, new strategies. It can be stimulating and provocative, leading to interesting new debates. Creativity does not mean absence of methodological rigour or ethical standards, but encourages full exploitation of the capacities of the human knowledge in conducting evaluations.

3.2 Evaluation Criteria

Evaluation uses a certain number of specific criteria to assess the relevance, performance and achievements of a project or programme. Contrary to the other methodological concepts and notions developed in the previous sections, the criteria are relatively standardised in the evaluation practice and did not lead to extensive literature and contradictory debates. There have been slight changes in the definition of some of them following the introduction and broader use of the Result Based Management model, insisting on the notion of performance management.¹² However, the international community, in particular through the OECD/DAC network, is working to harmonize the terminology in order to avoid possible confusion.

Relevance is the extent to which the objectives and/or project purposes of a project/programme remain valid and pertinent either as originally planned or as subsequently modified. The relevance of projects could be analyzed at different levels, for instance the relevance to the country's development strategy, to the target groups or beneficiaries' needs, to the IOM global mandate or to partners and donors policies.¹³

The analysis of the **performance** looks at the progress being made by the project or programme according to criteria, standards or performance indicators, and in relation to its overall objectives and/or project purposes. A restrictive assessment of performance could focus simply on the delivery and transformation of inputs into outputs, as presented in the project document, which would limit the scope and usefulness of the evaluation exercise. This is usually done in the activity or progress reports prepared by the programme managers during the life of the project.

In order to avoid such a restrictive approach, three criteria are commonly used to assess the performance of a project/programme. The **effectiveness** considers the extent to which a project or programme achieves its objectives/project purposes or

¹¹ For instance Michael Q. Patton as mentioned in Valéry Ridde's article, '*L'expérience d'une démarche pluraliste dans un pays en guerre: l'Afghanistan*', The Canadian Journal of Program Evaluation, Vol. 18 No 1, 2003 (ref. Michael Q. Patton, *Creative Evaluation*, 1987).

¹² 'Performance' and 'achievements' are central notions in the RBM model. This is less the case with the more traditional development assistance evaluation terminology, derived mainly from the logical framework approach, where focus is placed on effectiveness, both positive and negative impact as well as on relevance and usefulness. RBM also introduced the notion of 'outcome'. There is a tendency now to harmonize the terminology and work with both approaches on a complementary basis.

¹³ All the definitions in the section are from or adapted from the various papers on the glossary of evaluation terminology of the OECD/DAC. Sometimes, they have been adapted to IOM work. For the questions that can be related to each criterion and that can facilitate understanding, see Annex 3.

produces its desired results, taking into account their relative importance. It is equally important to examine, however, if changes would have occurred irregardless, without the implementation of the project or programme. Effectiveness is also used as an aggregate measure of (or judgment about) the merit or worth of an activity. **Efficiency** analyses how well resources in general (funds, expertise, time, etc.), or inputs, are used to undertake activities, and are converted to results. Sometimes, the definition also takes into account the analysis of alternatives. However, the notion of **cost-effectiveness** is more appropriate in that case as it mainly considers whether the objectives were achieved at minimal cost (or at the lowest possible cost), or whether the results or benefits justify the cost. Cost-effectiveness is different from a **cost-benefit analysis**, the definition of which varies, but which is centered on a monetary value and a comparison between costs and benefits.

'Achievements'¹⁴ includes criteria such as *impact* and *sustainability*. **Impact** is certainly the criterion raising major debates, especially due to the methodological constraints imposed by its measurement, as already pointed out under Section 2.1. According to OECD/DAC glossary, impacts are "positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended". The impact is sometimes defined under a more restrictive perspective, especially in the framework of the RBM approach that "limits the focus only to what is expressed as positive and expected achievements within the strategy, and further narrows down the focus to what is realistically attributable to the project".¹⁵ What is important to remember when discussing impact is first to properly define what is intended by 'impact' in the framework of the evaluation and then examine how it can be observed, measured and described. Due to the complexity of measuring it and the costs for conducting sophisticated research, the 'judgment of the wise' can sometimes be used as a solution for assessing globally the impact.

The RBM approach also introduced another notion used in project development and evaluation and that can create confusion with the notion of impact: '*outcome*'. **Outcome** can be simply defined as the achieved short-term and medium-term effects of an intervention's outputs.¹⁶ 'Result' and sometimes 'Impact' are used instead of outcome, but outcomes are more immediate and directly benefit target populations, whereas impacts are longer term with broader societal consequences or benefits. The **sustainability** is the durability of project's results, or the continuation of benefits once external support ceases. In IOM, this criterion is more appropriate for projects or programmes with capacity building, development or reintegration components than with resettlement activities or movements.

Other criteria, less often used, can still be mentioned here. The **validity of design** looks at the logic and coherence of the design of the project or programme. It is especially important with complex project or programme documents, but it is sometimes analyzed implicitly in the assessment of the relevance of the strategy of the

¹⁴ UNDP uses instead the notion of 'success' focusing mainly on the positive and expected achievements of the programme/project, in line with RBM approach.

¹⁵ OECD/DAC, '*Review of Current Terminology in Evaluation and Results Based Management*', December 1999.

¹⁶ According to what is called the 'Results Chain', the following sequence describes outcome: Inputs – Outputs – Outcomes – Impact. The definition can change however if, for instance, referring to the 'Outcome Mapping' methodology and behavioral change. In that case, outcomes are defined "as changes in behavior, relationships, activities, or actions of the people, groups, and organizations with whom a programme works directly" (See International Development Research Center publication on Outcome Mapping) – see Annex 1.

project or programme. The **unanticipated or unexpected effects** are positive or negative effects, but not foreseen when the activities started and not part of the objectives.

Four criteria have been specifically designed for the evaluation of humanitarian assistance in emergency.¹⁷ The analysis of **appropriateness** examines the extent to which humanitarian activities are tailored to local needs and to the requirements of ownership and accountability, as well as how the activities responded to changing demands in a volatile environment. **Coverage** is defined as the need to reach major population groups in life-threatening situations, wherever they are, with assistance proportionate to their needs, including protection. **Coherence** refers mainly to policy coherence, ensuring that all policies (security, trade, military, development etc.) take the humanitarian aspects and human rights into consideration. And **connectedness** is the need to assure that activities of a short-term emergency are implemented in a way that takes longer-term and interconnected approaches into account. OECD/DAC also considers important to examine the **co-ordination** issue, even if it can be analysed under 'effectiveness'. However, due to the complexity of emergency situations and the multiplicity of actors involved, it deserves sometimes to be specifically taken into account.

3.3 The Reference Data

All the criteria described above require a 'comparative basis' for conducting the analysis. For many of them, the expected **results**, as defined in the project or programme document, are an important reference, but not necessarily the only one. If a result can be achieved, it does not mean necessarily that the project has had the expected impact. This is why the logical framework methodology introduced the **indicators** to complement the references available, and to further clarify and more precisely define the objectives and project purposes and the expected impact on beneficiaries. According to OECD/DAC definition, an indicator is a "quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement, to reflect the changes connected to an intervention, or to help assess the performance of a development actor".¹⁸

The indicators will help to render more objective, subjective notions such as 'improve', 'reinforce', 'promote', 'comprehensive', 'enhanced', which are commonly used when writing objectives and project purposes. Definite words like 'install', 'increase from x to y', 'build', 'eradicate', 'number' are more appropriate for indicators. The indicators can be divided in sub-groups, for instance '*input indicators*' related to the inputs of a project, '*process indicators*' to the activities, '*outcome indicators*' to the project purposes or '*impact indicators*' to the goal or overall objective of the project. '*Indirect or proxy indicators*' are used when data directly related to an objective is difficult to collect for measuring the effectiveness and achievements. Indirect reference is used in replacement, for instance to measure the change of awareness in counter-trafficking activities through the increased number of TV debates around the subject. The

¹⁷ For further references: IOM, *Monitoring, Evaluation and Audit in Emergencies*, Ch. 7 EOM, 2002 available under the IOM Evaluation Webpage – technical references. See also OECD/DAC *Guidance for Evaluating Humanitarian Assistance in Complex Emergencies*, 1998; ALNAP, *Evaluating Humanitarian Action, an ALNAP Guidance Booklet*, 2003, www.alnap.org.

¹⁸ OECD/DAC, *Glossary of key terms in Evaluation and Results Based Management*.

introduction of the RBM model, and related performance measurement, added the notion of '*performance indicators*' to the terminology, which can be defined as measures describing how well a programme is achieving its objectives. The major difference with the general definition of indicators lies in the specific reference to planned objectives (and related planned results).

The indicators need in any case to be well defined, realistic, limited and measurable. If too ambitious for instance, they are not useful as it could become difficult to assess what is realistically achievable and attributable to the project or programme. As mentioned in the definition above, they can be quantitative or qualitative. Indicators are also used for monitoring the projects and programmes, and not only for evaluation. This is mainly why OIG is not formally tasked with providing technical assistance for the establishment of indicators, which is left to departments in charge of project development and monitoring. However, IOM field offices often contact OIG/Evaluation for guidance on indicators and, informally, OIG will continue offering its technical support when requested.

Two other notions will be briefly presented here: '*benchmark*' and '*baseline data*'. According to OECD/DAC glossary, **benchmark** is a "reference point or standard against which performance or achievements can be assessed. [...] A benchmark refers to the performance that has been achieved in the recent past by other comparable organizations, or what can be reasonably inferred to have been achieved in the circumstances". **Baseline data** describe existing conditions that provide a starting point against which progress can be assessed and comparison made. They are often detailed in the context or background section of a project or programme document, referring to the situation that will be addressed by the project or programme implementation.

“Evaluation has always been about learning, about how to be accountable, how to be transparent, how to learn from experience. Today’s questions therefore, do not evolve around whether it is desirable, but focus at whose learning we are talking about and, for what purposes.” (Paul G.H. Engel & Charlotte Carlsson)

4. UTILIZING, PLANNING AND PREPARING EVALUATIONS

Section 4 intends to focus on practical issues for IOM staff, especially regarding the use of evaluation reports and how to plan and prepare an evaluation. It will be less theoretical than the previous sections that deserve, however, to be part of the guidelines in order to have a better understanding of evaluation, a comprehensive view of the complexity of evaluation and of the basic requirements for a credible evaluation exercise, in line with expected quality standards.¹⁹

The last two sections will provide IOM staff with some technical guidance on the planning and preparation of evaluations. The planning will mainly consist in the inclusion of an evaluation component in the project or programme document during the development phase, and the preparation in the elaboration of Terms of Reference, the recruitment of consultants and the organization of their work. The guidelines will not focus on conducting evaluations as a vast majority of IOM staff will not themselves perform evaluations. A sample format for the report to be used by appointed consultants is available in Annex 4, and Annex 5 presents and describes the self-evaluation reporting.

4.1 Utilizing Evaluation: Accountability, Learning and Promotion

There are multiple understandings of evaluation utilization and focus will be placed here on what is called sometimes ‘instrumental use’. In recent years, the increased requests for evaluations arise mainly from the increased demand for **accountability**. This is especially true in IOM when Member States, as well as donors, were pressing IOM Administration to reinforce the evaluation function at the end of the 1990s. The projects and programmes have to increasingly justify their existence, the use of funds, their relevance in terms of strategy, their effectiveness in reaching objectives and their sustainability. In IOM, it is especially valid as the organization is still expanding its scope of work within the framework of its broad mandate, implementing new projects and proposing new strategies. Accountability is not limited to the notion of ‘control’ (or financial control), which is more specific to audit work, but is defined as an “obligation to demonstrate that work has been conducted in compliance with agreed rules and standards or to report fairly and accurately on performance results vis-à-vis mandated roles and/or plans”.²⁰ It also includes the proper use of resources, which is of importance due to the limited availability of funds allocated to development and humanitarian assistance and to the unlimited needs, including in the field of migration management.

¹⁹ Section 2 and section 3 also have a practical dimension but mainly for the few IOM officers in charge of the Evaluation function or for those conducting evaluations. Evaluations of programmes and projects in the field are usually performed by external consultants.

²⁰ See OECD/DAC glossary.

Accountability implies an agreement between two parties and the responsibility from one party to properly implement the projects and programmes funded. There is also an obligation to report back on the responsibility engaged, the report being for instance the evaluation itself. Contrary to activity or progress reports, often prepared by the programme managers themselves, evaluation reports are an in-depth study of the performance expected. A reward or a sanction can be applied based on the reporting. Concerning the 'reward' or 'sanction' in evaluation, a donor can decide to fund an extension of the project or programme based on the evaluation results, but contrary to what is often believed, the reward and sanction in evaluation are not personalized and should not be personalized. A failure in project implementation does not imply necessarily a weak management. External factors can interfere and become the main cause of failure. 'Personalized evaluations' are left to the staff performance appraisal mechanisms, the management of which falls under the responsibility of the human resources management functions. Accountability also requires transparency as already defined under Section 3.1.

A common definition of **learning** is the process by which individuals or organizations create, transfer and use knowledge in order to achieve positive change and realise their goals. The learning dimension of evaluation also led to the elaboration of different theories, approaches and perceptions of 'learning'. They will not be developed here and only the distinction between the '*developmental perspective*' and the '*knowledge perspective*' of evaluation and learning will be briefly presented.²¹

Under the developmental perspective, evaluation is conducted to improve institutional performance, the design of projects, programmes or strategies, institutional effectiveness and responsiveness, to assess the perception of the organisation among the beneficiaries, or to recommend changes in the organisation's activities. The knowledge perspective encourages, on its side, the acquisition of deeper understanding of the underlying factors of a problem, the relevance of the solutions proposed, or an increased understanding and knowledge of the theory and logic behind an action. In that case, evaluation is considered as having the potential to 'generate' knowledge.

Most of the evaluations conducted by OIG tend to use both perspectives in the way the findings, conclusions and recommendations are presented and used. The OIG also tries to encourage, in parallel to the learning dimension of the evaluation itself, the approach described as the 'Evaluative Thinking and Process Use' that should help "people in programmes learn to think and engage each other evaluatively. Teaching evaluative thinking can leave a more enduring impact from an evaluation than use of specific findings. [...] Those stakeholders actively involved in an evaluation develop an increased capacity to interpret evidence, draw conclusions, and make judgements".²² A prerequisite to thinking 'evaluatively' is, however, to train people in evaluation for using this new approach. This is not always possible because of the heavy workload

²¹ Another distinction is made for instance between the *improving-policy-feedback* approach focussing on policy-makers and donors learning, the *collective learning* approach involving all stakeholders in the learning process and the *societal learning* approach focussing on the societal learning and change.

²² See Michael Q. Patton, *On Evaluation Use : Evaluative Thinking and Process Use*, in the Evaluation Exchange, Vol. IX, No 4, Winter 2003/2004.

faced by most professionals in the field and the financial constraints in prioritizing and organizing such training.

One important question regarding learning is: can we really stop and take time to think, read and analyse evaluations as was the case 20 years ago, especially taking into account the incredible amount of reading, experiences exchanged and knowledge available through the effective and worldwide communication means, and the needs, in an 'accelerated' world, to move and respond more and more quickly to any demands? Everyone will agree that it is less and less possible and that there are rapid and efficient means for learning, in particular through e-learning, and through the new trends of learning by experience or learning by doing. The evaluator is no longer the only reference expert whose evaluations will be systematically used as a central input in programme development and policy decision.

This does not mean however that evaluations cannot be used for learning but they need to adapt to new realities. In order to promote learning, OIG's objective is to render evaluation more accessible, less exclusive and more attractive and interesting. Another initiative is to focus on the evaluation report structure and content. Key conditions for the evaluation reports to be attractive, and its findings and recommendations used, can be summarized as follow: they need to be interesting, professional, relevant, when possible innovative, and more importantly short. A maximum of 20 to 30 pages can be more easily digested than a detailed report of 100 pages, full of data and statistics, and can encourage the selection of the report among other available learning sources. It is evident that this cannot be applied to any type of evaluation but is possible with project and programme evaluations that are relatively simple covering a limited number of objectives (it will be more difficult to apply such a rule to evaluations assessing, for instance, the national education policy in a country).

It is also important to avoid being too confrontational with the report, to have the capacity to build confidence and to be ready to 'negotiate' the content for increasing the readiness to learn from the evaluation, without, however, denaturing the evaluation exercise and distorting the findings. A selective approach in the planning of evaluations to be conducted by field offices is also encouraged in the sense that innovative or sensitive projects or programmes may benefit more from the learning dimension of evaluation than traditional projects or programmes running smoothly. A mandatory request to have all projects and programmes evaluated does not help evaluation to be accepted as a learning tool, but encourages the perception of a 'threatening' control mechanism.

Three notions related to learning and used in the evaluation context must be mentioned: '*lessons learned*', '*best practices*' and '*recommendations*'. **Lessons learned** are generalizations based on evaluation experiences that abstract from specific circumstances to broader situations. They can be useful if used, which is not often the case for the simple reason that the commissioner and users are not necessarily interested in such generalisation or have no mandate to institutionalize it. **Best practices** when extracted from the same generalisation perspective suffer from the same 'uselessness'. In addition, best practices are selective and elitist by definition as they focus on the notion of 'best', ignoring in fact that 'worst practices' are as useful as best practices, if not more, for learning from a given experience. Focus for an effective and immediate utilisation is more on the **recommendations** of the

evaluations, which can be briefly defined as proposals aimed at enhancing the quality of a project/programme's design and implementation. Recommendations should be linked to conclusions (a **finding** is a 'factual statement' and a **conclusion** is a 'synthesis' of 'factual statements' corresponding to a 'specific circumstance').

A third use of evaluation is related to the **promotion** of the work carried out by an organization. An evaluation is providing an independent and objective view of the performance and achievements of a project, programme or policy. It offers an all-embracing and comprehensive view of the subject or object evaluated, and clarifies sometimes grey zones that were for instance previously questioned by donors or partners. The promotion of an activity, strategy or policy through evaluation will gain in credibility and will not be perceived as a pure marketing operation. The evaluation does not necessarily need to be positive and successful to be used as a promotional tool. An evaluation identifying weaknesses, proposing adjustments also shows the capacity of the organization to be open to criticism, transparent in its management, ready to learn from experiences and to adapt to changing needs. Too often unfortunately, staff members consider an evaluation highlighting weaknesses as potentially damageable to the organization.

4.2 Planning an Evaluation

As mentioned in the previous sections, the inclusion of evaluation in IOM project or programme documents is no longer mandatory and relies on the capacity of the Project Developers to properly assess the needs for an evaluation and to have clear ideas on the objectives and rationale of the evaluation. They have to examine the size of the project, the interest of the donor(s), partners and stakeholders in including an evaluation component, in particular in terms of accountability, the sensitivity and innovativeness of the project or programme, and its usefulness especially in learning from that experience, or as an objective basis for discussing possible extensions.

If they take the decision to include an evaluation component, the next step will be to **determine the type** of evaluation. As mentioned under Section 2.1, the most common type is a programme or project evaluation. It is important recalling here to avoid confusion by proposing impact evaluations, which are often complex and expensive, and participatory evaluations, which require the participation of stakeholders at every step of the evaluation exercise, rendering it also complex. Mid-term and final evaluations are the most common. As already mentioned, the selection depends of the duration of the project or programme. Ex-ante and ex-post are not frequent. The OIG/Evaluation can advise on the type of evaluation, if necessary.

The Project Developers should also avoid mentioning internal evaluations as internal evaluators are not easily available, mainly due to resource constraints, except for the self-evaluation to be conducted by the Project Manager. OIG conducts internal evaluations but the selection of projects or programmes to be evaluated rests at the OIG level. The Project Developers should consult with OIG before committing to any internal evaluation. Therefore, in most cases, the evaluations will be external, conducted by externally recruited consultants. Joint evaluations require the agreement of partners (for instance donors) and OIG should be consulted for appropriateness before proposing it. In both cases, a financial provision must be set aside in the budget for that purpose.

Depending on the type and complexity of the project or programme, the **budget provision** can vary, for instance if the project requires a lengthy evaluation exercise, if several field visits need to be organized or if detailed surveys have to be conducted for collecting sufficient and reliable data for the evaluation. As already mentioned surveys can be very expensive and are often limited to measuring the impact of a project or programme. Other methodologies for collecting data could suffice. The level of the overall budget of the project or programme is another important element to take into account. If the total budget is for instance USD 200.000, it will certainly be inappropriate to propose a budget of USD 50.000 for an external evaluation, except if the evaluation is part of a research and learning component of the project, which is, however, rarely the case.

The budget level is mainly established based on the cost of the consultant, which can vary from USD 100 to USD 2.000 per day of work, depending on the 'quality' and expertise of the consultant, and if they are locally or internationally recruited. If the project is of strategic importance for the field mission, or for IOM as a whole, and its budget high, the recruitment of an experienced and more expensive consultant may be planned. OIG considers, however, reasonable a fee for an international consultant ranging from USD 250 to USD 500 per day, not including travels and daily subsistence. Depending of the country, the fee for an experienced local consultant can vary from USD 90 to USD 300 per day of work. IOM very rarely uses teams of consultants due to the types of projects and programmes it develops.

The duration of the exercise has an impact on the costs too. For a complex project, the Project Developers could have to consider a minimum of 30 days of work, and for a less complex one the exercise could be finalized in 15 to 20 days. In summary, the financial provision for a three weeks evaluation exercise may be USD 6'000, without however considering surveys. OIG/Evaluation can also provide guidance on financial provisions and the necessity of conducting additional surveys. A budget revision can always be discussed with the donor(s) during implementation if changes occur in the planning of the evaluation, and additional surveys are, for instance, needed.

4.3 Preparing and Organizing an Evaluation

The preparation of the evaluation exercise starts with the elaboration of the Terms of Reference – ToRs. They can be prepared by the IOM Project Manager alone, or in full cooperation with the benefiting government, the donor(s) and/or other partners. In any case, it is important to have a broad agreement on the ToRs as they will be the basis of the evaluation exercise and more importantly, the formal reference for the donor(s), partners and for the consultant or team of consultants to be recruited. They also formalize the different ideas, objectives and discussions around the exercise, as well as the way to analyze and evaluate them. It is also possible to adjust the ToRs during the implementation of the evaluation when duly justified, but the evaluator cannot decide on it without the agreement of the Project Manager.

The ToRs should define a clear objective for the evaluation, explaining why the evaluation is undertaken and what the evaluation intends to accomplish, the scope of the evaluation or plan for conducting the evaluation, the issues to be addressed through the listing of the evaluation criteria that will be used, the products expected

from the evaluation, which may include in addition to the report itself particular strategies or recommendations on some specific areas, the methodology used, the party responsible for conducting the evaluation and finally the implementation arrangement, including a realistic timeframe and logistical arrangements.

All the points above can be included under the following sections of the ToRs. The **background** section summarizes the overall context and the project or programme that will be evaluated. The objectives and project purposes of the project or programme should be mentioned as they will be one of the main references of the evaluation. The overall and specific **objectives of the evaluation** are essential as they precisely frame the work and product expected from the evaluation exercise. It is important, however, to focus on key issues and prioritize. It is better to have an evaluation that examines a few issues in depth rather than one that looks into a broad range of points superficially. A clear description of the evaluation criteria that will be covered is also useful. The section on the **methodology** explains the methodologies that will be used during the evaluation. It is also important to be rather precise, especially if there is a need to conduct surveys that might have an impact on the duration of the evaluation exercise or on the costs for the evaluation, when not already foreseen in the budget. Finally, the section on the **resources and timing** will specify the costs for the evaluation and who will cover them, the tasks of the consultants if an evaluation team is selected, the logistical arrangements, the work plan and the deadlines for submission of the first draft and final reports. A sample format for the ToRs covering all the points above more in details is available in Annex 6.

The next step is the recruitment of the consultant(s). There is no central evaluation roster managed by OIG as the Office does not have enough external evaluations conducted regularly to justify its update. OIG can provide some references on a case-by-case basis. There are other solutions available and one of them is to consult the donor(s) in the field if they would like to propose a consultant, which can be seen by the donor(s) as a guarantee of the independence and objectivity of the evaluation. Major donors often have lists of consultants ready to work in the country. Another solution could be to issue a call for proposals open to international and national consultants. That process is, however, more time consuming than the identification through informal contacts and it will be more difficult to have objective recommendations on their professional capacities. Finally, universities can be approached locally as professors are sometimes ready to conduct evaluations as local consultants, in addition to their work.

It is evident that the most important aspect of the recruitment is to guarantee that the candidate(s) really has an appropriate background and specific experience. This could be problematic but interviews can be conducted before the final selection. As soon as the candidate is selected, a formal contract can be signed using the IOM consultant contract form, but the ToRs must be attached as an annex. The IOM office responsible for the project has the responsibility to properly organize the evaluation exercise, including the interviews with key partners, and provide all information requested by the consultant(s), and other information deemed necessary for the evaluation.

The IOM office has the possibility to provide its input to the evaluation exercise through the comments on the draft report to be submitted by the consultant(s) before finalizing it. A report can be considered satisfactory if the spirit of the ToRs is fully met and if the

content is clear, easy to understand and logically well justified. If the Project Manager has major disagreements with the content of a report, OIG can be consulted for advice on how to proceed to avoid paying a consultant and not publishing the final product. OIG/Evaluation can also provide feedback on the quality of the report in terms of methodology.

As soon as finalized, the report can be distributed. A debriefing can be organized with the participation of the donor(s), partners and stakeholders during which the consultant can present the report and key findings and recommendations. There is no restriction in the distribution policy to the public except for IOM transport services and resettlement programmes organized under a contractual basis with a government. In that case, the external dissemination should be agreed upon with the donor. OIG should be kept informed of the issuance of any evaluation report as all reports are posted in the IOM Evaluation Webpage. Concerning the utilization of the report, Section 4.1 already covered the issue.

ANNEX 1: BIBLIOGRAPHICAL REFERENCES

The references are not exhaustive and only cover the main material used for preparing the guidelines. OIG can provide more detailed bibliographical references on evaluation, as well as relevant websites.

The quotes on the top of each section are from the following sources:

- Section 1: Michael Quinn Patton, *Qualitative Methods and Analysis for Development Evaluation*, IPDET Training, Ottawa, 2001
- Section 2: Ray C. Rist, *On Evaluation Utilization: from Studies to Streams*, in 'The Evaluation Exchange – Reflecting on the Past and Future of Evaluation', Vol. IX, No. 4, Winter 2003/2004, Harvard Family Research Project - HFRP
- Section 3: Eleanor Chelimsky, *The coming Transformations in Evaluation*, in 'Evaluation for the 21st Century, a Handbook', E. Chelimsky & W. R. Shadish, 1997
- Section 4: Paul G.H. Engel, Charlotte Carlsson, *Enhancing learning through evaluation: Approaches, dilemmas and some possible ways forward*, Paper presented at the European Evaluation Society Conference, Seville, 2002

Other references:

- OECD/Development Assistance Committee (DAC), *Principles for Evaluation of Development Assistance*, OECD/GD(91)208, 1991
- OECD/Public Management Committee, *Best Practice Guidelines for Evaluation*, PUMA(98)3, 1998
- OECD/DAC, *Guidance for Evaluating Humanitarian Assistance in Complex Emergencies*, DCD/DAC/EV(99)3, 1998
- OECD/DAC, *Promoting Practical Sustainability*, by AusAID, 33rd Meeting, 2000
- OECD/DAC, *Results Based Management in Donor Agencies*, by A. Binnendijk, 34th Meeting, 2001
- OECD/DAC, *Review of the Quality of Evaluations – The 'Minimum-Sufficient' Group of Evaluation Standards*, 2nd Meeting, 2004
- OECD/DAC, *Review of Current Terminology in Evaluation and Results Based Management*, DCD/DAC/EV(2000)2, 1999
- OECD/DAC, *Glossary of Evaluation and Results Based Management (RBM) Terms*, DCD/DAC/EV(2000)6, 2000
- OECD/DAC, *Glossary of Evaluation and Results Based Management Terms, Terms and Definitions Submitted by Members*, 33rd Meeting, 2000
- OECD/DAC, *Glossaire des Termes Utilisés dans le Domaine de l'Evaluation de l'Aide au Développement*, par les Ministère des Affaires Etrangères et Ministère de l'Economie, des Finances et de l'Industrie, France, 33^e Réunion, 2000
- OECD/DAC, *Glossary of Evaluation and Results Based Management Terms*, DCD/DAC/EV(2001)3, 2001
- OECD/DAC, *Glossary of Key Terms in Evaluation and Results Based Management*, with French and Spanish translations, 2002
- E. Chelimsky & W. R. Shadish, *Evaluation for the 21st Century, a Handbook*, Sage, Thousand Oaks, 1997
- J. Z. Kusek, R.C. Rist, *Ten Steps to a Results Based Monitoring and Evaluation System, a Handbook for Development Practitioners*, The World Bank, 2004

- R. Picciotto, E. Wiesner, *Evaluation and Development, The Institutional Dimension*, The World Bank, 1998
- ALNAP, *Evaluating Humanitarian Action, an Alnap Guidance Booklet*, ODI London, 2003
- ALNAP, *Alnap Review of Humanitarian Action in 2003, Field Level Learning*, ODI London, 2004
- M. Q. Patton, *Organizational Development and Evaluation*, in The Canadian Journal of Program Evaluation, Special Issue, 1999
- Sida, *Looking Back, Moving Forward: Sida Evaluation Manual*, Sida, 2004
- M. Q. Patton, *Enhancing the Quality and Credibility of Qualitative Analysis*, in Health Services Research 34:5, Part II, 1999
- Harvard Family Research Project, *Where We've Been and Where We're Going: Experts Reflect and Look Ahead*, in 'The Evaluation Exchange – Reflecting on the Past and Future of Evaluation', Vol. IX, No. 4, Winter 2003/2004
- M. Bamberger, *The Evaluation of International Development Programs: A View from the Front*, in American Journal of Evaluation, Vol.21, N01, 2000
- J. Z. Kusek, R.C. Rist, *Building a Performance-Based Monitoring and Evaluation System: the Challenges Facing Developing Countries*, The World Bank, 2001
- American Evaluation Association, *The Programme Evaluation Standards, Summary of the Standards*, in AEA Website <http://www.eval.org>, 1994
- American Evaluation Association, *Guiding Principles for Evaluators*, in AEA Website <http://www.eval.org>, 2004
- UNDESA, *Impact Evaluation: Are We Looking in the Wrong Direction for the Promised Land?*, UN IAWG on Evaluation, discussion paper, 2000
- The World Bank, *2002 Annual Report on Evaluation Capacity Development*, The World Bank, 2003
- V. Ridde, *L'Expérience d'une Démarche Pluraliste dans un Pays en Guerre: L'Afghanistan*, in the Canadian Journal of Program Evaluation, Vol.18, No 1, 2003
- E. J. Davidson, M. Scriven, *Journal of Multi Disciplinary Evaluation*, Number 1, 2004
- S. Taurelli, *From Evaluation to Benchmarking: How Evaluated Organizations can Learn*, paper presented at the EES Conference, 2000
- M. Q. Patton, *On Evaluation Use: Evaluative Thinking and Process Use*, in The Evaluation Exchange, Vol.IX, No 4, Winter 2003/2004
- W. Meier, *Results-Based Management: Lessons in Implementation*, Results-Based Management Group, 1999
- H. B. Weiss, *Democratic Evaluation*, The Evaluation Exchange, Vol. XI No 3, Harvard Family Research Project, 2005

- The World Bank, *Monitoring and Evaluation: some Tools, Methods and Approaches*, Operations Evaluation Department – OED, 2002
- International Fund for Agricultural Development - IFAD, *A Guide for Project M&E*, in <http://www.ifad.org/evaluatio>, 2002
- UNDP, *Results Oriented Monitoring and Evaluation*, UNDP, 1997
- USAID, *Preparing an Evaluation Scope of Work*, in Performance Monitoring and Evaluation Tips, No 3, 1996
- USAID, *Using Rapid Appraisal Methods*, in Performance Monitoring and Evaluation Tips, No 5, 1996
- USAID, *Selecting Performance Indicators*, in Performance Monitoring and Evaluation Tips, No 6, 1996
- USAID, *Guidelines for Indicator and Data Quality*, in Performance Monitoring and Evaluation Tips, No 12, 1998

- The World Bank, *Designing Project Monitoring and Evaluation*, in Lessons & Practices, The World Bank – OED, No 8, 1996
- UNEG, *Standards For Evaluation in the UN System*, United Nations Evaluation Group, 2005
- UNEG, *Norms for Evaluation in the UN System*, United Nations Evaluation Group, 2005

- IOM, *An Evaluation Strategy for IOM*, Office of Programme Evaluation, 1998
- IOM, *Evaluation Guidelines Including how to Perform Self-Evaluations*, Office of Programme Evaluation, 1998 (rev. 1, 2001)
- IOM, *Research and Evaluation Methodology for Mass Information Activities*, MMS/EVA, 1999
- IOM, *Functions of the Office of the Inspector General*, GB 1259, 2000
- IOM, *Lessons Learned from Evaluation Activities: IOM's Response for Follow-up to UN GA Resolution A/56/201*, OIG, 2002
- IOM, *Monitoring, Evaluation and Audit in Emergencies*, OIG/EPC – Emergency Operations Manual, 2002
- IOM, *Accountability and Learning in Evaluation*, IOM Chief of Mission Training, OIG, 2002
- IOM, *Project Development Training - Evaluation*, OIG, 2002
- IOM, *IOM Project Development Handbook*, MMS, 2005

ANNEX 2 : OTHER TYPES OF EVALUATION

- ***Strategic Evaluation:*** An evaluation of a particular issue aiming to advance a deeper understanding of the issue, reduce the range of uncertainties associated with the different options for addressing it, and help to reach an acceptable working agreement among the parties concerned. It is usually adapted when urgency of the issue poses high risks to stakeholders and has generated conflicting views.
- ***System-wide Evaluation:*** An evaluation used in emergency situation, which covers the response by the whole system to a particular disaster or emergency.
- ***Partial System Evaluation:*** An evaluation also used in emergency, which covers only a part of the system. It can be related to thematic or sector evaluations.
- ***Single Agency Response Evaluation:*** Still in emergency, an evaluation which covers the overall response by a particular agency.
- ***Single Agency/Single Project Evaluation:*** An evaluation which covers a single project undertaken by a single agency in emergency situation.
- ***Real Time Evaluation:*** An evaluation implemented in emergency situation aiming to provide a rapid feedback on humanitarian operations and to be an immediate catalyst for improvements in organizational and operational performance. The methodology cannot be rigid and flexibility and adaptability are required, although guaranteeing quality.
- ***Sector Evaluation:*** An evaluation of a variety of aid actions all of which are located in the same sector, either in one country or cross-country. A sector covers a specific area of activities such as health, industry, education, transport or agriculture.
- ***Cluster Evaluation:*** An evaluation which analyses a set of related activities, projects or programmes to identify common threads and themes.
- ***Empowerment Evaluation:*** An evaluation promoting close involvement between the evaluator and the project/programme participants, so as to produce more meaningful and useful evaluation results. Empowerment evaluation is necessarily a collaborative group activity, not an individual pursuit.
- ***Stakeholder Evaluation:*** An evaluation which involves agencies, organizations, groups or individuals who have a direct or indirect interest in the development assistance, or who affect or are positively or negatively affected by the implementation and outcome of it. Stakeholders work together to develop and finalize instruments and procedures, produce recommendations, and make decisions throughout the evaluation process (related term: *Participatory Evaluation*, which focuses on methodology).

- ***Quasi-experimental Impact Evaluation:*** An evaluation which compares different groups before and after programme implementation in order to assess the programme impact and value added of further investments. It uses rapid and economical studies that combine exploitation of existing data sets with rapid sample surveys, tracer studies, interviews, etc.
- ***Meta Evaluation:*** An evaluation of an evaluation, performed to judge its quality, merit, worth or significance. An aggregation of a series of evaluations of the same kind or of findings is a not a meta-evaluation but a *meta-analysis of evaluations* and might not be evaluative.
- ***Cross-section Evaluation:*** A systematic evaluation of various evaluation reports on a specific project type, on projects involving one particular sector, or on one particular instrument or theme, designed to review and possibly update existing development policy directives.
- ***Incorporated/built-in Evaluation:*** An approach to implementation that involves fairly continuous self-evaluation by principal actors and participants, according to pre-established criteria related to the purpose and goal of the assistance.
- ***In-depth Evaluation:*** An approach which consists of focusing evaluation or a part of an evaluation precisely on a category of outputs, or on a group or category of impacts.
- ***Theory Based Evaluation:*** An evaluation which focuses on an in-depth understanding of the workings of a programme or activity, the 'programme theory' or 'logic'. It needs not assume simple linear cause-and-effect relationships but maps out the determining or causal factors judged important for success, and how they might interact.
- ***Democratic Evaluation:*** An evaluation approach which addresses critical evaluation issues such as dealing with power relations among stakeholders, including stakeholders' perspectives, and providing useful information to programmes. Power redistribution is accomplished by 'democratizing knowledge' and holding all groups, including the client, mutually accountable.

ANNEX 3
SOME QUESTIONS RELATED TO EVALUATION CRITERIA

- 1) **RELEVANCE:** make the project/programme more responding to problems to be solved to what extent the project/programme continues to make sense in light of current reality, and to what extent its objective and project purposes remain valid and pertinent.
- Is the project/programme consistent with external reality (which can include living conditions of the target group(s), national and regional government policies, administrative capacities or institutional and cultural factors)? Have Gender mainstreaming issues been taken into account?
 - What were the references used for assessing the “reality” and how objective are they? What does support this assessment, e.g., government documents, media reports, donor mission reporting, UN documents?
 - Is the project/programme implementation strategy consistent with stated objectives and IOM priorities and mandate? Is it complementary to other agencies and government activities?
 - Are the global context and the problems to be solved still the same (unless the project/programme managed to solve them)?
 - Are there any elements which could lead to the conclusion that the overall objective or (one of) the project purposes have to be dropped, revised, amended or new ones added in order to make the project/programme more in line with the current reality and needs?
 - To which extent does it already become evident that additional or complementary activities/projects need to be implemented to?
- 2) **VALIDITY OF DESIGN:** to what extent the project/programme is properly designed taking into account the economic, social, political and developmental context and the problems to be solved.
- Is the project/programme document designed according to IOM project development guidelines and logical framework approach, including indicators, budget and relevant information on monitoring and evaluation?
 - Are the problems clearly stated? Do the overall objectives, project purposes and activities properly address the problems to be solved? Is there any major incoherence in the logical development of the project/programme?
 - Do the expected results mentioned in the project/programme document cover the project purposes? Does the project/programme document list all results (which are one of the main references for measuring achievements together with indicators of achievement when listed)?
 - Was the context of the project/programme sufficiently developed in order to avoid confusions? Does the project/programme make reference to complementary activities/projects which, if not mentioned, could lead to the conclusion that the global design was not properly apprehended?

3) **EFFECTIVENESS:** to what extent the project/programme produced desired global output and outcome, and more specifically, to what extent the project purposes produced the desired results.

- What are the project/programme results and indicators that can be used as a reference? Is there sufficient information, data or material available to verify the achieved results and/or indicators?
- Is the project/programme effective in reaching planned results, in bringing planned changes? Does a major gap exist between planned results and achieved results (indicators listed and verified data)? How can it be interpreted? Is it due to inappropriate design only, or are there other elements which can explain that gap?
- Were the activities sufficiently well defined and implemented in order to reach relevant results? If results are not reached, would other activities have resulted in the results being reached?
- Did the achieved results reach the target group as planned? Did the achieved results reach any indirect beneficiaries and was it planned as such or is it an unexpected consequence?
- Are the achieved results more numerous than the planned ones? What else can be done to make the project/programme more effective?
- Would the recorded changes have occurred even without the implementation of the project/programme? To what extent are identified changes the result of the project/programme rather than extraneous factors?

4) **EFFICIENCY:** how well resources in general (funds, expertise, time, etc.) or inputs are used to undertake activities, and are converted to results.

- What measures have been taken to ensure that resources are efficiently used?
- Are the project/programme expenditures reports and records sufficiently well prepared to guarantee transparency and to make a global analysis of the costs incurred to reach the results obtained?
- Are the project/programme expenditures in line with the agreed upon budget, and the costs incurred consistent with the strategy of the project?
- Did the funds expended give the possibility to reach the expected results?
- Are there additional results, which were not expected, achieved without additional costs? If additional costs were charged to reach those results, was the investment level appropriate and useful?

5) **COST-EFFECTIVENESS:** whether the objectives were achieved at minimal cost (or at the lowest possible cost), or whether the results or benefits justify the cost.

- Were the costs proportionate to the results achieved? Was the investment worthwhile for the achieved results?
- Could the activities have been implemented with fewer resources without reducing the quality and quantity of the results? Could a different type of project/programme or strategy have solved the same problems and reached the same results at lower costs?
- If the conclusion is that not all cost levels were fully justified, which ones were too high, staff costs, office costs, operational costs? Should the resources have been used for a more worthwhile purpose related to the project/programme?

- 6) **IMPACT:** how activities of the project/programme contributed to a change in a situation, intended or unintended, positive or negative, and what the project/programme was expected to bring.
- Is the project/programme document sufficiently well designed to identify which impact was expected from the project/programme and attributable to it?
 - As actually measuring a change in the situation may be complex and expensive, how can you identify and measure it in a way that can be verified? Are there sufficient observable elements, evidences and/or data which could lead to a firm conclusion about the impact of the project/programme and exclude alternative explanations?
 - Either positive or negative, does the impact come from the project/programme activities, from external factors or from both? What would have occurred without the project/programme?
 - What do the target groups and other stakeholders perceive on themselves to be the impact of the project/programme?
 - Does the impact only concern the target population or is another population affected by the activities of the project/programme? What can be observed in terms of gender balance and impact?
 - Were there any impacts that were not foreseen by the project/programme? Is there a possibility to draw conclusions, in addition to the impact on the target group, on a global impact at the social level, political level, economic level, or on institutional capacity?
- 7) **SUSTAINABILITY:** to what extent the project/programme benefits continue after external support is no longer available, or to what extent the project/programme continues functioning without external financial support.
- Do the overall objective and project purposes contain an element of sustainability? Do appropriate indicators and means of verification exist in order to verify it?
 - Is the project/programme supported by local institutions and well integrated with local social and cultural conditions?
 - Did local partners and stakeholders participate in the planning and implementation of the project/programme to guarantee sense of ownership and interest in the sustainability?
 - Do partners have the financial capacity to maintain the benefits of the project/programme on a longer-term basis? Is the technology utilized appropriate to local conditions?
 - Are the results obtained sufficient to draw any conclusion on the sustainability before the actual end of the project/programme? Are some activities already on-going without anymore support from the project/programme?
 - What should be done in order to guarantee the sustainability, if necessary?
- 8) **UNANTICIPATED EFFECTS:** any unforeseen effects, positive or negative, that resulted from the project/programme.
- Is it possible to observe a result or an effect, positive or negative, which has been indirectly and unpredictably "created" by the project/programme?

- Are these effects complementary to other results or are they completely independent? Did these unanticipated effects affect (positively or negatively) other results?
- If negative, is it possible to correct them in order to avoid such an effect?
- If positive, is it possible to build other activities on these effects which can benefit the project/programme and for which a revision of the strategy is needed?

9) **CAUSALITY:** to what extent some factors or events affected the results of the project/programme.

- Are there any external events or factors which strongly affected the results obtained versus the results planned? Is there a possibility to correct the situation, if needed, in the framework of the project/programme?
- Are there evidences that the event or factor reported really affected the results or are there some other factors, also internal to the project/programme, which play a role in the deviation?

ANNEX 4 : EVALUATION REPORT FORMAT

Annex 4 is mainly intended for appointed evaluators/consultants inquiring about the IOM evaluation report format. It can, however, also be used by any IOM staff interested in conducting internal evaluations and not wishing to follow the self-evaluation formats presented in Annex 5.

There are various options for writing an evaluation report, depending on many factors such as the methodology used or agencies' requirements. Findings, conclusions and recommendations could, for instance be put in evidence under specific sections or, as in the case of IOM, they could be presented as they appear in the logical flow of the report.

OIG considers that, in order to complete a good evaluation report, the following are required:

- The content of the report should be clear, statements well justified and the logic behind the evaluation exercise respected.
- As IOM projects and programmes are not too complex that they require very detailed analysis, the main text of the report should not exceed 30 pages. This makes for easier reading than a detailed 100-page report. Annexes containing detailed information may be appended if deemed appropriate.
- The report should be interesting to read and bring new elements/perspectives to the subject evaluated. An evaluation report is not an activity or progress report.

It is evident that, in addition to the three conditions mentioned above, the evaluation norms and standards set out in the guidelines should also be used in writing the report.

The logic of an evaluation exercise is in fact at the basis of the report format and can be summarised as follows:

- 1) The evaluation exercise is intended to assess a project, programme, activity, theme, policy or organizational structure, which are also influenced by external factors. All aspects relating to effective performance need, therefore, to be correctly appraised and presented and this can be done in a section setting out the topic under evaluation, its background and context.
- 2) The quality and validity of an evaluation also depends on properly selected methodology. A section should, therefore, be included briefly setting out the objectives of the evaluation, methodologies selected, constraints in applying them or deviations from the Terms of Reference, if any.
- 3) Before making any analysis, it is important to state the problems, describe their context and list the findings as empirical data and evidence. Specific sections will then treat each of the various issues under evaluation. It is necessary to avoid statements that are dealt with in the analysis or judgment as this could lead to confusion. These sections must only deal in facts.
- 4) As soon as all the information needed for the evaluation is collected, consolidated and presented, the analysis may commence using the different evaluation criteria: relevance, effectiveness, efficiency etc. The conclusions and recommendations for each of the criteria selected can be mentioned under related sections, flowing logically from the analysis.

Other sections could be added if, for instance, generalization or lessons learned can be extracted from the analysis as shown in the example below.

It is important also to include an *'Executive Summary'*, not exceeding three pages, as well as a section entitled, *'Summary of Conclusions and Recommendations'*. It is useful to have a global view of the results of the evaluation particularly as there is so often a lack of time for the actual report to be read in detail. Annex 1 usually includes the ToRs and Annex 2 the bibliography, or documentation reviewed, as well as the interviews conducted. Other annexes can be added to the report but should be selected so as not to overload it.

The Table of Contents below illustrates the format suggested in the four points above. The impact is not mentioned as the project was still on-going. Other examples of OIG-conducted evaluations can be found in the IOM Website, under the Evaluation Webpage:

http://www.iom.int/en/who/main_structures_evaluation.shtml#internal

EXECUTIVE SUMMARY

1. INTRODUCTION

- 1.1 The MIDA Concept and *MIDA Italy*
- 1.2 Evaluation Scope and Methodology

2. THE MAIN COMPONENTS OF *MIDA ITALY*

- 2.1 Research, Database and Promotional Activities
- 2.2 Institutional Agreements and Main Partners
- 2.3 Investing in Development

3. RELEVANCE OF *MIDA ITALY*

- 3.1 Relevance of the *MIDA Italy* Strategy
- 3.2 Relevance of the Countries and Target Groups Selected

4. PROJECT PERFORMANCE

- 4.1 Effectiveness of Preparatory Phase
- 4.2 Effectiveness in Skills Transfer and Development Investment
 - 4.2.1 Ghana
 - 4.2.2 Ethiopia
- 4.3 Overall Management by IOM Rome (*NB. It includes efficiency and sustainability analysis*)

5. SOME LESSONS LEARNT FOR APPLICATION TO MIDA

- 5.1 MIDA Ownership
- 5.2 Promotion of MIDA and Practical Experiments
- 5.3 MIDA: A Broad Concept
- 5.4 MIDA Management

6. SUMMARY OF CONCLUSIONS AND RECOMMENDATIONS

ANNEX 5 : SELF-EVALUATION

Why a 'self-evaluation'?

Self-evaluation is sometimes described as an important feature of learning organizations and critical to personal and professional growth, if implemented in full 'self-transparency' and with a strong 'self-learning' intent. Its major weakness lies in the guarantee of objectivity of the exercise, the self-evaluation being prepared and conducted by project/programme management itself.

However, by keeping the process simple (compare to a fully-fledged evaluation), not too time-consuming and with specific reporting requirements, the self-evaluation can produce a wealth of useful information on projects/programmes and can be beneficial in the longer term for project/programme planning. It is also important to take into account that project/programme managers do not have necessarily an expertise in evaluation. If not well designed, the system could then result in inconsistent reports or in too favorable evaluations.

Self-evaluation is also a cost-effective exercise and certainly the least expensive solution compare to independent evaluations. Although the self-evaluation does not imply any consultant fees (either internal or external), the project/programme manager must ensure that all costs are budgeted. They can include for instance travels to the field, or costs for a short assessment/survey to be conducted locally before the evaluation exercise.

As for any evaluation, particular emphasis needs to be placed on the objectives of the self-evaluation and what is expected from the exercise. If the only purpose is to do it because it is required in the project document, it is likely that the result will reflect that thinking, and the self-learning dimension and quality will be strongly affected. It is important to consider how the self-evaluation can benefit the project/programme as a whole, the project/programme management and IOM colleagues working in the project/programme or on similar projects/programmes in other countries.

A step by step process

When preparing a self-evaluation and in order to properly write the self-evaluation report and related matrices, the project/programme manager should raise some preliminary and preparatory questions and follow precise steps:

- 1) What is important to know about the project? Is it accomplishing what it was set up to do, or not? What are/were the problems? What can be done differently? Who should be approached during the self-evaluation exercise? Etc.
- 2) One way to answer some of those questions is to collect all the relevant data available on the project from the project document, the progress reports, from regular project monitoring, or from correspondence with external partners. It is recommended to record all resources used in the evaluation and their sources in an annex.
- 3) The next step consists of sorting out data in order to complete the different boxes of the matrices (see below), mainly those covering **effectiveness**, starting with the most concrete elements, in particular the results achieved and data related to indicators. If needed, more

information should be collected through further research in project documentation, or through other specific activities, such as interviews. Developing a list of standard questions to be asked of all interviewees could facilitate a formal search for data. It is important to talk to a representative number of persons, rather than just one or two, so as to avoid any hidden biases. If the information required for completing the matrices is not available from internal documents or interviews, this may be a signal of some gaps in the project's data collection activities.

- 4) Other evaluation activities can also be considered, for example focus groups or surveys. As those activities may require specialized skills, such as for leading focus groups and drawing conclusions from them in an organized way, consultants may have to be sought locally. Related costs would have to be borne by the project budget.
- 5) The following step will consist of assessing the other evaluation criteria, mainly **relevance, efficiency, cost-effectiveness, impact and sustainability**, using also the questions listed under Annex 3. For those elements where insufficient concrete information exists to allow an objective assessment, more information should be collected as discussed above, for example interviewing persons who have some knowledge of the project, but who have no stake in its outcome.
- 6) When sufficient information allowing an objective assessment is gathered, the matrices can be completed. It is important to focus on main issues and to summarize them even if the information has to be somewhat abbreviated.
- 7) The last step will be to write the report, as described below. The report can be distributed internally or externally. This is left to the decision of the project/programme manager and Chief of Mission. The self-evaluation report can be added as an annex to the donor interim/final report.

The self-evaluation report

Every project/programme manager certainly agrees that the more simple and practical the self-evaluation reporting system, the more efficient the process. This is the reason why the core of the self-evaluation exercise has been placed on matrices. Only some additional considerations will have to be written up as the narrative part of the report, as mentioned below.

Two formats are proposed below: **Format A** for standard projects/programme self-evaluations and **Format B** for emergency activities with a shorter narrative part and including notions specific to humanitarian assistance evaluation (for instance connectedness or coverage). If the project/programme manager prefers to use other reporting formats, for instance the full format proposed under Annex 4, this is possible but the exercise could end up being more constraining in terms of logical development and reporting requirements for meeting evaluation standards. OIG/Evaluation can provide examples of self-evaluation reports.

The self-evaluation report should reflect an evaluative assessment of the situation of the project based on pertinent and concrete information, and not simply give project management's opinions. Precise information, either quantitative or qualitative, gives the reader the possibility to better understand and accept the validity of the self-evaluation report. To state that "IOM provided effective support to the government and migration management remains a key priority for the government" should be a summary statement that clearly and logically comes from

factual elements included in the matrices, and not a statement that could appear as unjustified opinion only.

The project/programme manager preparing the report should also take into account that the reader does not necessarily have much knowledge about the project or programme being evaluated. It is therefore of particular concern that the report is clear for all readers.

The narrative part of the report should be prepared after completion of the matrices, even if presented before them, as they contain most of the information necessary for the assessment of the performance and achievements. The section should, however, not repeat in detail what has been already stated in the matrices and should not exceed four pages. The narrative part should be used:

- To provide the readers with a context which may help them to better understand the content of the report, in particular of the matrices;
- To present the outputs of the matrices in a more sequential, summarized and/or general way;
- To highlight issues that are important but not found in the matrices and that may be weighed against other elements, e.g. political considerations.
- To elaborate on specific important issues underlined in the matrices;

Some basic definitions must be recalled here: a *finding* is a 'factual statement'; a *conclusion* is a 'synthesis' of 'factual statements' corresponding to a 'specific circumstance'; a *recommendation* is a 'prescription' on what should be done in a 'specific circumstance'; a *lesson learned* is a 'generalization' which does not refer to a specific circumstance but to a 'type' of situation.

FORMAT A

The narrative part:

1. **EVALUATION RATIONALE:** the section should summarize the *objective(s) of the evaluation*, as well as the key issues to be addressed by the evaluation.
2. **THE PROJECT AND ITS CONTEXT (or BACKGROUND):** the section should briefly provide the reader with the following information, much of which may be available in the project document:
 - The context in which the project has been developed, starting with a brief description of the institutional, social, political and economic contexts, as well as any pertinent consideration on gender issues. A brief description of the stakeholders is also recommended, as well as any complementary activities or projects, either implemented by IOM or by another agent.
 - The problems that the project seeks to address, and its duration. The project document should list such information, mainly through the overall objective and project purposes. The overall objective and project purposes as presented in the project document can be recalled under this section.

- The strategy of the project, the linkages among objectives, activities and expected results. Considerations can be brought on the indicators to be used and on the implementation and management arrangements. It can include information on IOM partners, the government(s), the donor(s), international agencies and the NGOs associated, as well as on external partners who played an indirect role in the project implementation.
- Finally, the comparative advantages of IOM for the execution of the project can be recalled under this sub-section, if appropriate.

3. MAJOR EVALUATION FINDINGS AND CONCLUSIONS: the section should highlight the findings that are particularly noteworthy, those that may require further explanation and major conclusions regarding achievements, for instance:

- The contribution of the project to meeting the needs of the target population (if other groups not mentioned as target population benefited from the project, they should be mentioned under this section);
- any impact of the project on beneficiaries beyond the expected one, i.e. economic, social, political impact;
- factors that affected the implementation of the project, for instance government commitment or participation of the stakeholders;
- indication of eventual success.

4. SUMMARY OF RECOMMENDATIONS: what are the main recommendations proposed in the matrices and drawn from project/programme's implementation. They are recalled under this section for easy reference, but they need to be worded in the same way as in the matrices to avoid confusion.

The matrices:

The self-evaluation exercise is not intended to be an in-depth evaluation of the project, even if information and data collected could be used for such a purpose. The information contained in the matrices should therefore be summarized, the need for details supporting overall statements being balanced against the need for brevity.

The matrices include recommendations with information on the person or office responsible for follow-up. It is certainly expected that recommendations of a self-evaluation will be directed towards improvement of the project/programme. However, if recommendations are directed towards a Department at Headquarters or an MRF, the self-evaluation should be shared with concerned entities and attention brought to the recommendation.

Matrix 1: Effectiveness of the project – analysis through project results and indicators: the matrix is based on observable facts, assessing the indicators and results associated to the overall objective and project purposes against their deviation versus planned results or indicators. List any recommendations that can be drawn about project effectiveness.

Matrix 2: Other evaluation criteria – general analysis and recommendations. The matrix includes an analysis of the relevance, efficiency, cost-effectiveness, impact and sustainability of the project. Considerations on the validity of design, unanticipated effects or causality could be

added here, if deemed necessary. When some remarks are specific to one project purpose only, it should be mentioned as such in order to avoid confusion. If the project/programme manager wants to keep separate the comments according to the overall objective and project purposes, she/he can do so. Finally, confusion should be avoided between effectiveness (Matrix 1) and efficiency (Matrix 2). Further references to these criteria can be found in the guidelines.

MATRIX 1: EFFECTIVENESS OF THE PROJECT

Analysis through project results and indicators

- A. Overall Objective:** Mention it here (from project document). In most cases no specific results are listed for measuring achievements of the overall objective. Indicators related to it can be used as reference. If not available from the project/programme document, indicators can be established for the self-evaluation exercise. Normally, only one overall objective is mentioned per project/programme. If more than one, add a new table for each objective.

Indicators [I]	Achievements related to Indicators [I]	Analysis of Effectiveness: analysis of progress towards achievements	Recommendations
[I]: List any indicators	[I]: List corresponding reality, or factual observations based on the verification of indicators (findings)	Analyze the level of achievement of the overall objective: summarize to what extent the project produced desired global outcome through analysis of any deviation from indicators (conclusions).	List any recommendations or corrective action. Propose follow up responsibilities if necessary.

- B. Project Purpose 1:** Mention the first project purpose (from project document). For project purposes, results are listed. If, in addition to results, indicators are mentioned, add them to the results in the first column.

Results [R] & Indicators [I]	Achievements related to results (achieved results) and to indicators	Analysis of Effectiveness: analysis of progress towards achievements	Recommendations
As per project document, list all planned (expected) results and/or corresponding indicators.	List corresponding reality, or factual observations based on the verification of achieved results and/or indicators (findings)	Summarize to what extent the project produced the desired results (specify each result for clarity, R1, R2 ...). Provide comments on deviation(s) versus planned results and/or indicators – positive or negative (explain <u>briefly</u> reasons and consequences if appropriate) (conclusions).	List any recommendations (or corrective action) related to the effectiveness (Propose follow up responsibilities if necessary).
[R1] +[I1]:	[R1] +[I1]:	[R1]:
[R2] + [I2]:	[R2] +[I2]:	[R2]:
[R3].....	

B. Project Purpose 2: Mention the second project purpose (from project document) and prepare the same table. Continue the same exercise for Project Purpose 3, Project Purpose 4 ...

Results [R] & Indicators [I]	Achievements related to results (achieved results) and to indicators	Analysis of Effectiveness: analysis of progress towards achievements	Recommendations
<i>As per project document, list all planned (expected) results and/or corresponding indicators.</i>	<i>List corresponding reality, or factual observations based on the verification of achieved results and/or indicators (findings).</i>	<i>Summarize to what extent the project produced the desired results (specify each result for clarity, R1, R2 ...). Provide comments on deviation(s) versus planned results and/or indicators – positive or negative (explain briefly reasons and consequences if appropriate) (conclusions).</i>	<i>List any recommendations (or corrective action) related to the effectiveness (Propose follow up responsibilities if necessary).</i>
[R1] +[I1]:	[R1] +[I1]:	[R1]:
[R2] +	[R2]	[R2]:

MATRIX 2 : OTHER EVALUATION CRITERIA

General Analysis and Recommendations

CRITERIA (please refer to the definitions in the core text)	RECOMMENDATION (if any)
Relevance: <i>briefly analyse the relevance of the project/programme. As mentioned in the guidelines, it can be analyzed under various perspectives: the relevance of the strategy, the relevance to IOM mandate, the relevance of target groups selected etc. Both findings and conclusions should be considered here.</i>	<i>List any recommendations (or corrective action). Propose follow up responsibilities if necessary.</i>
Efficiency and Cost-effectiveness: <i>briefly analyze both criteria (findings and conclusions).</i>	<i>Idem as above</i>
Impact: <i>briefly analyse the criteria taking however into account the constraints for measuring impact – see the guidelines for further reference. Outcome analysis could be considered. In addition to the findings and conclusions, an explanation on the limits for measuring impact should be added, if necessary.</i>	<i>Idem as above</i>
Sustainability : <i>briefly analyse the criteria (findings and conclusions)</i>	<i>Idem as above</i>

Optional criteria: see the definitions in the guidelines

CRITERIA	RECOMMENDATION (if any)
Validity of design: <i>can also be discussed under 'Relevance'</i>	<i>Idem as above</i>
Causality: <i>briefly analyse the criteria (findings and conclusions)</i>	<i>Idem as above</i>
Unanticipated effects: <i>briefly analyse the criteria (findings and conclusions)</i>	<i>Idem as above</i>

FORMAT B:
(For Emergencies)

SELF EVALUATION REPORT

Project or Activity: *Title of the project or brief explanation of the activity: for instance Humanitarian Evacuation.*

Project Document Available: **Yes** **No**

Starting Date: **Estimated Duration:**

Total Budget:

Budget Spent: **Remaining Funds:**

Context: *Give a brief description of the overall context, IOM overall strategy and response, main actors and collaborating partners.*

Constraints: *For instance accessibility to the affected zones, security situation, identification of beneficiaries, protection measures in the management of camps or for transportation. Concerning the protection, this issue can also fall under the analysis of effectiveness if it is part of IOM tasks to guarantee the protection and human rights of the assisted populations.*

Summary of Conclusions and Recommendations: *Based on the analysis below.*

ANALYSIS

1) Relevance: *Three levels of analysis are usually pertinent for IOM under this criterion: 1) the relevance of IOM strategy/assistance to the situation on the ground (the notion of 'coherence' explained in the guidelines can be analysed here); 2) the relevance of the intervention to the needs of the beneficiaries (this is rather similar to the notion of 'appropriateness' explained in the guidelines); 3) relevance of the targeted population which is very similar to the notion of 'coverage' also mentioned in the guidelines.*

Conclusions and Recommendations: *List the conclusions and recommendations on the relevance analysis.*

2) Efficiency and Cost-Effectiveness: *Both concepts can be analysed under the same section as it is not expected to go into the financial analysis in any depth. Efficiency can briefly assess the financial management tools in place and the allocation of funds between operations and staff and office costs. Cost-effectiveness can also be briefly analysed, for instance the cost per passenger with company x could be compared with company y, or the cost per temporary shelter could be analysed according to costs incurred by other partners.*

Conclusions and Recommendations: *List the conclusions and recommendations for the efficiency and cost-effectiveness analysis.*

3) Connectedness: *This is rather important for IOM in emergency situations due to active and close collaboration with other international partners and various countries.*

Conclusions and Recommendations: *List the conclusions and recommendations for Connectedness.*

N.B. *The other criteria presented in the guidelines can be analysed here (they can include timeliness, impact or sustainability). However, due to the short timeframe of an emergency assistance, the analysis of the impact and sustainability are often not appropriate or feasible.*

4) Effectiveness: *The analysis is made for each objective/project purpose/activity of the IOM assistance. If a project document or SPROUT is available, the objectives/project purposes, planned results and indicators must be listed according to the document. If only contractual agreements or MoUs are available, the main activities or sectors of activity will be listed with corresponding results or 'reasonable' expected results if not specifically mentioned (Indicators of the 'Sphere' project can also be used as reference).*

Objective 1 (or Project Purpose 1, or Activity 1): *For instance 'To organise humanitarian evacuation of civilians from conflict zone'.*

Planned/expected results, indicators	Achieved results at the time of the evaluation	Analysis of the Effectiveness	Conclusions or recommendations
1) Targeted population identified and registered	1) Database has been created and a team has started registration of the population. 50% of the caseload registered	1) No major problem in the identification of population to be assisted. Minimal data needed and recorded for the operation. Timeframe in line with established work plan. Good coordination with other partners	1) Proceed as planned. No need to increase registration speed. If possible, to include additional data on the database
2) Transport by air to safe locations organized.	2) A charter company contracted with at least 5 planes available and countries of destination clearly identified.	2) Average of 4 flights per day organized. Problems have been encountered for allocating a country of destination to the refugees. No problems reported at the reception of the beneficiaries.	2) No need for the time being to increase the number of flights per day. However, need to better organize the attribution of a country of destination according to more precise criteria, e.g. family reunification.
3) Internal transportation from the camp to the airport organized and working	3) 10 buses rented for the internal movements. Also used for transportation between the different camps	3) No specific problems. Access to buses properly monitored. Sometimes overloaded with personal belongings.	3) Consideration should be given to rent trucks for transportation of belongings.

Objective 2 (or Project Purpose 2, or Activity 2): *To organise the medical screening of the displaced persons before evacuation.*

Planned/expected results, indicators	Achieved results at the time of the evaluation	Analysis of the Effectiveness	Conclusions or recommendations
1) <i>International and national doctors recruited through MHD/Hqs and working</i> 2) ...	1) <i>One international and three national doctors recruited and working.</i> 2) ...	1) <i>Problems have been encountered in performing the work due to the high number of persons evacuated every day.</i> ...	1) <i>Organize as soon as possible additional recruitment – 2 nationals at least.</i> ...

Objective 3:

ANNEX 6TERMS OF REFERENCE FORMAT

As mentioned in the guidelines, the Terms of Reference (ToRs) are an important piece of the evaluation exercise as they clarify what are the objectives, the methodologies that will be used, the composition of the team of evaluators, as well as the resources and timing for the evaluation. They will constitute the main formal reference, along with a contractual agreement between all partners concerned and with the evaluator(s) on what is expected from the evaluation.

Expanding adequate time and efforts in preparing good ToRs is a guarantee of the evaluation's quality, relevance and usefulness. Some general questions should be asked when preparing the ToRs, or after having prepared the first draft, in addition to the specific questions related to the evaluation criteria that will be presented under the relevant sections of the ToRs. These questions make up a short 'checklist' which is useful before finalization, for instance:

- Can all of the evaluation's objectives and questions mentioned in the ToRs be objectively covered and analyzed?
- Are there alternatives listed in the ToRs to cover questions during the exercise that may not be answerable (for instance on the impact)?
- Are the expectations regarding the evaluation exercise and reporting requirements properly explained?
- Are there possibilities for using alternative methodologies during the evaluation or should the ToRs be rigid in that respect?
- Are time allocated and resources sufficient to conduct the evaluation?
- Is the expertise required for the consultant(s) adequate and realistic (for instance compared to consultant's costs on the international market for a given expertise)?
- Do the ToRs guarantee transparency of the exercise, for instance not focusing only on what is perceived as a positive outcome?
- Is IOM support to be received by the consultant(s) throughout the exercise properly identified?

The ToRs can be developed in great detail (**Format A**), which is especially useful for complex projects/programmes and evaluations, or they can focus on key issues, which is the most common format in IOM practice (**Format B**). A mix of both formats can be considered if focus needs to be placed on some of the sections of the ToRs, for instance the description of the work of a team of consultants.

FORMAT A:**1. BACKGROUND OF THE PROJECT/PROGRAMME:**

The section is a detailed description of the context of the project/programme and of the project/programme itself. The objective(s) and project purposes of the project/programme must be listed as they are the main reference to measure/evaluate the relevance, performance, achievements, outcome and impact. Some of the expected results and indicators may be mentioned as they are also an important reference for measuring the performance of the project/programme. This is especially valid if, for instance, they are not clearly established in the project/programme document.

The section can also make reference to other initiatives that could complement or interfere with the project/programme and that are important to raise with the evaluator. It can give a brief description of the role of each partner and stakeholder in the project/programme implementation.

2. OVERALL OBJECTIVE OF THE EVALUATION:

The Section can briefly describe the reasons for carrying out the evaluation, the intended audience and use, the stakeholders involvement in the evaluation exercise, as well as the type of evaluation selected (mid-term, final, project/programme, thematic...). It can briefly state, for example, if more focus should be placed on a single activity, on a set of activities, on the strategy for achieving an objective, or if the evaluator should avoid focussing on a specific aspect at the detriment of a more global analysis of the performance.

The Section will then summarize all the points above by precisely stating the overall objective of the evaluation: 'To evaluate...'. The overall objective should, however, remain general as a Section will cover the specific objectives of the evaluation exercise.

3. SPECIFIC OBJECTIVES OF THE EVALUATION:

The Section will highlight the issues to be studied in detail. The information can be divided in sub-sections according to the evaluation criteria, and for each criterion a series of specific objectives can be mentioned. For instance:

Relevance:

- *To assess the relevance of the overall strategy of the project;*
- *To examine the relevance of the target group selected with respect to the needs of the affected population etc...*

Effectiveness:

- *To evaluate to which extent the results achieved are in line with expected outcome;*
- *To analyse the effectiveness of the collaboration between main actors in meeting the expected results...*

Efficiency: etc.

Instead of formulating specific objectives, questions can be used for each criterion. For instance instead of 'To assess the relevance of the overall strategy...', the following question can be asked: 'Is the project strategy clearly explained and in line with Government policy?'. Sample questions related to the evaluation criteria can be found in Annex 3.

Whatever method is selected, it is important to frame the statements in a way that they are understandable and 'analysable' with empirical evidence. Vague questions or objectives lead to vague answers and analysis. It is also important to limit the objectives or questions to key issues and management priorities. If necessary, comments or foot notes can be added to the objectives or questions to assist the evaluator(s) in the analysis (for instance, if papers were already developed on the subject covered by the question).

The Section can also specify what kind of recommendations and lessons the evaluation should provide. This is especially valid for complex projects or programmes that could lead to a great number of recommendations. By mentioning which recommendations should be made in priority, the evaluator can better frame its analysis. For instance, the organisation can be interested in all conclusions that can be reached regarding the effectiveness but would like to focus the recommendations related to its management capacity only.

4. METHODOLOGY OF THE EVALUATION:

The Section puts in evidence which methodologies will be used for conducting the evaluation. The main methodologies are described in the guidelines: documentation review, field visits, interviews and surveys.

It is important to go into detail in the description for complex evaluations especially if there is a need for additional surveys to collect data and to measure the achievements and impact, or to establish a clear plan for a field visit explaining, for example, why some locations are not selected. Information on the elaboration of an interview guide could also be mentioned in the section, together with the responsibilities for organizing interviews and field visits.

Details on the information sources can also be very useful for some complex evaluations to facilitate the work of the evaluator(s) or to focus on some specific aspects, such as gender analysis.

If there is a need for preparatory work, for instance discussions with stakeholders for refining the ToRs or for a proper selection of participatory techniques, that should also be mentioned here.

5. REPORTING:

The Section will specify the reports that should be delivered: inception report, end-of- field visit notes for the file from the evaluator(s), draft report(s) and final report. End-of-field visit notes for the file can be requested, for instance, with a team of evaluators, each travelling in a different country. They will facilitate the consolidation of data but can also be used as justifications in the analysis. The Section can mention if a formal presentation of the draft report by the evaluator(s) should be organized.

The Section can also specify the language, dates of delivery, including dates for commenting the draft report(s), number of copies, distribution responsibility, costs for printing, format layout, the type of annexes, or maximum number of pages.

6. EVALUATION TEAM:

The Section is especially useful with a team of evaluators in order to clearly specify the size of the team, the expertise and qualifications required, language skills, gender requirements, who will act as a team leader, what will be covered by each evaluator and how the team will work together for consolidating the report.

If only one evaluator is recruited, then the section can briefly describe the expertise required and qualifications, including language skills.

7. RESOURCES AND TIMING:

The Section should highlight how much the evaluation will cost, who will cover the costs, as well as other resources needed for conducting the evaluation, for instance in terms of logistical arrangements for field visits. The costs should include not only the evaluator(s) fares, but also additional costs such as surveys, interpreters or secretarial services. It should also specify how payments will be made to the evaluator(s) and under which conditions, in particular regarding the quality of the reports.

A precise work plan should also be mentioned including the timing for recruiting the evaluator(s), for organizing the field visits, the dates for submission of the different reports, for their distribution and for formal presentation of the evaluation results.

FORMAT B:

As mentioned above, Format B is simpler and is the most common in IOM. What should be specified in each section will not be repeated here, but only what is dropped compare to Format A in order to simplify the elaboration of the ToRs. Some Sections are dropped or merged.

1. BACKGROUND OF THE PROJECT/PROGRAMME:

No need to go too much in detail in the presentation of the project/programme. The summary in the beginning of the project/programme document is sufficient, together with the overall objective and project purposes as they give a good picture of the expectations.

2. OBJECTIVES OF THE EVALUATION:

Sections 2 and 3 can be merged. The Section can briefly describe the reasons for carrying out the evaluation and formulate the overall objective for the evaluation. The specific objectives or the evaluation questions can be mentioned under the same section. No need to specify all the

evaluation criteria; the statements made (either in the form of an objective or of a question) can include what the related criteria are and what is expected in the analysis.

The point remains, however, valid that vague questions or objectives lead to vague answers and analysis. It is also important to limit the objectives or questions to key issues and management priorities.

3. METHODOLOGY OF THE EVALUATION:

All the information mentioned under Format A should be maintained here, but with less detail. Section 4 on the reporting requirement can be merged under the same section. If the evaluation requires the recruitment of one consultant only, some information can also be included here. However, if a team of evaluators is planned, the specific section should be maintained.

4. RESOURCES AND TIMING:

The same information as in Format A can be mentioned.