

Evaluation Use in the UN System

Conclusions from the Data

UNEG S02 Group

This document summarizes the efforts made by the Working Group on Strategic Objective 2 (S02) – UN entities and partners use evaluation in support of accountability and program supporting case studies.

Forward

The United Nations Evaluation Group (UNEG) in its 2014 – 2019 strategy identified the use of evaluation as one of the key strategic objectives and sets as a goal that “UN entities and partners use evaluation in support of accountability and programme learning”. The working group under this strategic objective has gathered data from within the UN system and contrasted these findings with academic research from outside the UN system. This document, “Evaluation use in the UN system” presents the results of this effort.

UNEG is highly committed to the efforts of enhancing the use of evaluation and fostering the adoption of evidence-based policies, particularly within the international development community. By identifying underlying factors that contribute to an enforced evaluation use, this study aims to improve evidence based decision-making throughout the UN system and to contribute to a greater effectiveness of development actions.

The use of evaluation is a theoretical concern not only for evaluators and evaluation specialists but also for the entire community of development actors. It has been argued that the effective use of evaluation findings should represent the main objective of an evaluation exercise. The study identifies six messages that provide the UN evaluation community with a set of practical solutions to implement in their evaluative work. These findings are complemented by case studies that provide the reader with practical examples on how use of evaluation has been increased in specific situations across the UN system.

I would like to take this opportunity to invite all UN entities to strengthen the use of evaluation in their respective organizations and to thereby contribute to accountability and program learning within the UN system. Finally, I would like to express my very great appreciation to the ‘Working Group on Strategic Objective 2’, for their valuable work in bringing evaluation into the forefront of the UN system.

Marco Segone
Chair
United Nations Evaluation Group

Forward

Evaluation use remains a perennial issue in evaluation. It has been the subject of discussion and debate for decades in the evaluation literature, and there is now considerable understanding of the different types of possible uses of evaluation findings. There is also considerable advice in the literature on what can or should be done to enhance use. Nevertheless, evaluation findings and recommendations often continue to be reported as being under utilized. At the same time, use is often seen as a key aspect of quality evaluation. Clearly, this remains an important issue of research for the evaluation profession.

This UNEG study on use of evaluation therefore is a welcomed addition to this needed research. The UN system is a significant producer of evaluations around the world and enhancing use in the UN system is a needed focus of UN evaluators, managers and decision makers. This study is the first in a programme of research on evaluation use in the UN system to be undertaken over the next years. UN Members should actively support these efforts which aim at making UN programmes more effective in contributing to the new SDGs and being more accountable to Member states.

This UNEG study has a number of good features. It builds on the large literature on evaluation use and links it to the UN context. It uses surveys of UN agencies, structured interviews with a range of stakeholders and case studies to assess the state of evaluation use. It also explores the mechanisms at work that encourage or discourage use of evaluation findings, essentially providing insights into the theory of change for evaluation use. It summarizes its findings into six, easy to understand, Key Messages.

This initial report is sensibly used to outline the next steps in the UNEG research programme on evaluation use. The research questions and discussion points identified, as well as the details of the case studies, are all worthy of exploring and provide a good basis for the future research. The result should be quite helpful in understanding how to enhance use in the UN system and valuable additions to the broader issue of evaluation use generally. Given that the study does find systematic use of evaluation in the UN to be weak, future research needs to explore and build up what specific practices can be adopted to enhance different types of use. The current study identifies quite a few issues that can be usefully researched. And given the range of different organizations in UNEG, future research could usefully explore how specific organizational contexts affect evaluation use, a topic often not well addressed in the literature.

I look forward to future findings and contributions to the literature on this critical issue in evaluation from this valuable UNEG initiative.

John Mayne
Independent Evaluator

Preface

Table of Contents

Forward	2
Forward	3
Preface.....	4
Table of Contents	5
Executive Summary.....	6
Introduction	7
Objectives and Methodology.....	7
Results	8
The six key messages	9
Use.....	9
Our Key Messages	12
Key Message 1	12
Key Message 2	18
Key Message 3	22
Key Message 4	26
Key Message 5	30
Key Message 6	33
Next Steps & Future Research.....	36
Discussion points	36
Research questions:	37
Bibliography.....	39
Annex 1. Research Methodology.....	44
Objectives	44
Data.....	44
Research Questions	45
Survey instrument	46
Annex 2. Use of Evaluations – Theory and Data from the UN System.....	49
1. Levels of use.....	49
2. Defining evaluation use.....	50
3. Different purposes of evaluation.....	51
4. Users and types of use	52
5. Evaluation in the UN system	54
6. Use factors.....	54
Annex 3 Detailed Conceptual Framework.....	56
Annex 4 - Case Studies on Successful Use of Evaluation.....	58

Executive Summary

This research was carried out by the United Nations Evaluation Group (UNEG) Working Group on Strategic Objective 2 (SO2), in order to better understand how UN agencies use evaluation, and to identify the factors that support or hinder evaluation use. Towards these objectives, four distinct data gathering exercises were undertaken. This document presents the findings, key messages and supporting case studies from this research.

The Working Group found that reported levels of evaluation use vary within and across agencies; evaluation use does not appear to be systematic across the board. Many respondents shared anecdotal evidence of use, but did not consider that use in their agencies was high overall.

The main results extracted from the data are presented as six 'key messages'. These confirm the relevance of factors already identified in the literature to the UN system, and point to important linkages between them. The key messages are:

1. Users and stakeholders should be involved and consulted throughout the evaluation process.
2. The support of senior decision-makers is key, as is their commitment to the implementation of the recommendations.
3. Evaluators need to ensure that recommendations are feasible and relevant.
4. Independent evaluations must attempt to capture organizational realities.
5. Management responses and follow-up processes must take place and be adequately supported.
6. The sharing of findings enables cross-organizational learning and use.

Three main questions came out of the research:

- I. What are the specific mechanisms that make the key messages so important? Are they context-specific? Do they always hold true?
- II. What effective strategies have various evaluation offices or UN agencies used to strengthen their practice around each of the six key messages?
- III. How can organizations be encouraged to take action to strengthen practices and processes that support evaluation use?

The intention is to focus the 2015-16 work plan of the SO2 Working Group on these three questions, with a clear view to improving evaluation use in UN agencies.

Introduction

As part of its 2014 - 2019 Strategy, the United Nations Evaluation Group (UNEG) defined four strategic objectives: 1) Evaluation functions and products of UN entities meet the UNEG Norms and Standards for Evaluation; 2) UN entities and partners use evaluation in support of accountability and programme learning; 3) Evaluation informs UN system-wide initiatives and emerging demands; and 4) UNEG benefits from and contributes to an enhanced global evaluation profession¹.

This document summarizes the efforts of the Working Group on Strategic Objective 2 (SO2) in the year 2014/15. Specifically, it presents the findings and key messages of our research into evaluation use in the UN system, with supporting case studies.

Objectives and Methodology

The goals of this working group were to better understand how UN agencies use evaluation and to identify the factors that support or hinder evaluation use. There is considerable literature on evaluation use, and we were interested in finding out whether the same factors are at play in the UN system as in other settings. The results of this research will feed into the process of defining our work on SO2 for 2015/16.

Towards these objectives, four data gathering exercises were undertaken:

- I. A literature review of the most relevant pieces of academic and organizational writing about evaluation use;
- II. An online survey of UN evaluation users and practitioners, as well as external evaluation practitioners;
- III. Semi-structured interviews with UN evaluation users and practitioners; and
- IV. Preparation of case studies where evaluation was useful and used.

To keep our survey instruments simple, we created three categories of evaluation users:

(1) Practitioners

- Evaluators / UNEG members
- Evaluation community/associations
- Partnership and network organizations (ECG, IOCE, EvalPartners etc.)

(2) Immediate Users

- Planners
- Programme staff

¹ <http://uneval.org/document/detail/1459>

- National governments / collaborating partners
- (3) Intermediate Users
- Management
 - Donors and other resource partners
 - Governing body
 - NGOs and similar organizations

While helpful for basic analysis, this typology was too broad for some research needs. For example, our data does not differentiate between respondents from management or from the governing body. Different groups within the three categories might have had very different experiences with evaluation, but we were unable to capture those differences with our main survey and questionnaire instruments.

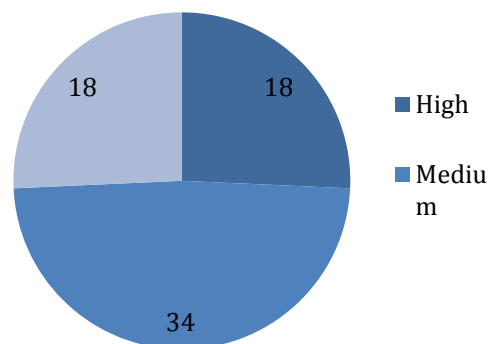
More information on the research methodology can be found in Annex 1.

Results

In terms of the levels of reported use within and across agencies, our findings are in line with the wider evaluation literature. Around a half of survey respondents rated evaluation use in the UN system as ‘medium’, which we defined as ‘periodic use’. A quarter considered use to be ‘high’ (‘systematic use for decision making, new project cycles, corrective actions’) and another quarter ‘low’ (‘infrequent use’).

However, different categories of users responded differently. 40% of ‘immediate users’ reported use to be high, compared to around a quarter of all respondents, and only 12% of intermediate users. We cannot determine at this stage whether this relates more to differences in perception, or to actual levels of use across agencies. However we can conclude that evaluation use is not systematic. Many respondents shared anecdotal evidence of use, but did not think that use was very high overall. Many thought that evaluation was most often used when it supported actions that were already planned by decision-makers.

Reported Use Level - Survey



There was a general perception that more could be done to generate greater and better evaluation use. However, no instances were found of evaluation misuse, and most respondents agreed that their agency, and the wider UN, are getting better at ensuring that evaluations are used.

The key messages

Based on the data we extracted six key messages, specific to the UN context. These largely concur with factors identified in the wider literature, and point to important linkages between them. However, some specificities of the UN system, for example the governance structure, make a difference to evaluation use patterns.

The key messages are for the most part simple and uncontroversial. Nonetheless, our data showed that they are still far from being universally applied. As such we believe it is worth re-emphasizing the importance of the messages and spreading the word, with the intention of improving use practices in the UN system.

The six key messages

1. Users and stakeholders should be involved and consulted throughout the evaluation process.
2. The support of senior decision-makers is key, as is their commitment to the implementation of recommendations.
3. Evaluators need to ensure recommendations are feasible and relevant.
4. Independent evaluations must attempt to capture organizational realities.
5. Management responses and follow-up processes must take place and be adequately supported.
6. Sharing of findings enables cross-organizational learning and use.

In this way we hope to identify some priority areas to improve evaluation use in the UN system, and for further research. To this end, for each of the key messages identified, we propose some points for discussion and future research.

Use

All evaluations are intended to be used; anything else would be absurd. Indeed, Patton (1997, in Herbert 2014, p.389) claims that, regardless of their quality, evaluations that are not used tend to be considered failures. This research does not intend to argue that use is the *most* important factor in judging the success of an evaluation, but simply that it *is* important, and more should be done to enable it.

Types of use

The literature defines many types of evaluation use, and there is debate over the terms ‘use’ and ‘utilization’. More detail about this is presented in Annex 2. For this report, we choose to refer to evaluation ‘use’, and focus on three types (with greatest emphasis on the first):

- I. **Instrumental** (or direct) **use** - the adoption and implementation of an evaluation recommendation;
- II. **Conceptual use** –the evolving conceptualization and understanding of those who design strategies and programmes; and
- III. **Process use** - changes to implementation or programming resulting from interactions between the evaluation team and key stakeholders.

The literature also distinguishes between two different *purposes* of evaluations - ‘accountability-driven’ and ‘learning-driven’ (Laubli Loud and Mayne 2014, p. 6). Overall, learning-driven evaluations are considered more likely to engender instrumental and conceptual use (de Laat in Laubli Loud and Mayne; Hawkins in Laubli Loud and Mayne 2014; Patton 2008), and are given greater emphasis in this research. However, our survey shows that accountability is still the main driver of evaluations in most agencies, with over 60% of respondents agreeing that the biggest factor in commissioning evaluations was the need for evidence on results or performance.

Use factors

In the literature we found over 100 factors affecting evaluation use, which we grouped together for simplicity, as summarized in Table 1 below. We isolated eight individual factors, grouped into two categories: factors related to the context of the evaluation, and factors related to the evaluation activities. These eight factors formed the basis of the conceptual framework for the data collection undertaken for this research. More information on these factors is available in Annex 3.

Table 1: Summary of use factors

Context in which the evaluation takes place	
	Evaluation culture
	Organizational structure and incentives
	Characteristics specific to the evaluation activity
Evaluation activities	
	Evaluation design and process
	User/stakeholder involvement
	Evaluator qualities
	Evaluation product
	Post-evaluation process

We found that our respondents did not tend to give all of these factors equal relevance. For example, user/ stakeholder involvement was very strongly stressed by our respondents, whereas organizational structure and incentives were barely mentioned.

Our **Key Messages**

This section describes the six ‘key messages’ emerging from this research. These are the factors considered most meaningful and relevant for supporting evaluation use in the United Nations system. The messages relate to various use factors, each relevant at different stages of an evaluation process, but all necessary to ensure effective use. Together they make up a chain of actions, processes and circumstances that must be in place to ensure that an evaluation is used.

Here we present the messages and relate them to the factors and mechanisms predominant in evaluation literature. For each key message, we provide a visual representation and a written description of how it supports evaluation use, alongside supporting evidence from our data. **We do not attempt to represent the full landscape of factors influencing use.** Rather, we focus on components considered by our respondents to have greatest potential to promote evaluation use in the UN system.

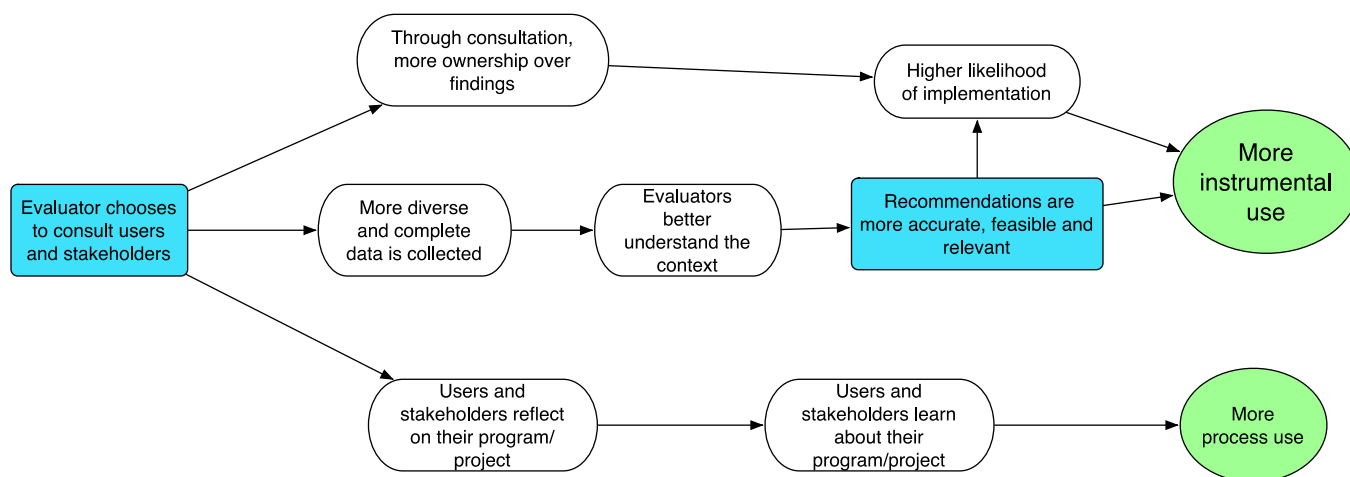
Key Message 1

Users and stakeholders should be involved and consulted throughout the evaluation process

Evaluators have long known about the importance of involving users and stakeholders in all aspects of the evaluation process, and the message seems to have been widely accepted. The literature points to several benefits to this: it means that stakeholders develop more ownership over the findings; it improves the quality of the recommendations; and it encourages process use.

Mechanisms at play

Figure 1: Mechanisms linking stakeholder engagement and evaluation use



I. When consulted, users and stakeholders develop more ownership over findings.

Most stakeholders develop a greater sense of ownership over evaluation findings when involved in defining the evaluation activities, asked to contribute data and information, and engaged in discussing the findings. This is especially important for the intended users of the evaluation findings, as the more ownership they feel, the more they are likely to buy-in to and implement the recommendations of the evaluators.

II. The consultation process improves the quality of the recommendations by enabling evaluators to collect better data and better understand the context.

Involving stakeholders is necessary to collect good information about the projects, programmes, policies or organizations being evaluated. It is a necessary step for evaluators to better understand the programme context, to come to valid conclusions and make useful recommendations.

It is recommended that stakeholders are involved from the stage of developing the terms of reference (TORs), to ensure that evaluation questions are relevant and useful, and that the data collection plan includes all available monitoring data. This gives evaluators better insights and ultimately ensures that recommendations are more accurate, feasible and relevant. As a result, recommendations are more likely to be adopted.

III. Consultation and engagement increases process use.

Although not a key focus of this report so far, process use is an important aspect of evaluation use, and depends on stakeholders being involved in the evaluation process. In this way, lessons can be shared with them along the way, enabling knowledge to be transferred throughout the evaluation.

Process use is more likely to happen when users and stakeholders can reflect on their experience as part of the evaluation process, and learn from their engagement in the evaluation. One of our interviewees confirmed:

“The final report often does not come as a big bang, it has been the many moments prior to the actual finalization of the report that have brought about change. Therefore: the engagement strategy prior to and during the evaluation process may present precious insights.”

Another advantage of stakeholder involvement in the evaluation process is that it exposes them to evaluation methods. This can help stakeholders to better understand and apply evaluative approaches as part of a more ‘scientific approach’ to programme design and internal performance assessment.

Data

Our survey presented an encouraging picture of stakeholder involvement across the UN system: 90% of UN respondents to our survey rated ‘stakeholder involvement’ as an ‘important’ or ‘very important’ factor for use²; approximately two-thirds stated that stakeholders are systematically engaged in the evaluation process; and a similar number considered that stakeholders are usually systematically made aware of the evaluation process.

This suggests that, while stakeholder involvement in evaluations is not yet universal in the UN system, most agree that it should be. This message was also strongly conveyed by evaluation practitioners in the open-ended questions about how to improve the use of evaluation, with several mentioning the importance of engaging and empowering stakeholders, increasing their awareness of evaluation findings.

The same message was reflected in our interviews. Most interviewees agreed that the evaluation should be participatory at all stages and that knowledge transfer should take

² The full survey questions can be seen in Annex 1: Research Methodology

place throughout an evaluation. However, many instances were reported of findings being shared with intended users and important stakeholders only at the end of an evaluation, either in the final report or at the dissemination workshop. One interviewee deplored that “It’s a failure of the evaluation if participation only takes place at the end”. Interviewees reported this lack of stakeholder involvement to be more common for accountability-driven evaluations, even though stakeholder engagement is considered important for all types of evaluation including accountability driven evaluations, to ensure that conclusions and recommendations are valid.

Despite the mixed picture currently, most interviewees considered that stakeholder involvement is becoming more widespread in UN evaluations. For example, the Independent Evaluation Office at the Global Environmental Facility is currently reforming its evaluation processes to ensure more consistent stakeholder engagement.

Notes from the literature

As noted above, the literature largely promotes the systematic involvement of stakeholders at all stages of the evaluation process. One of the main supporters of this is Michael Quinn Patton, whose lifelong work on user-focused evaluation proposes a highly participatory, stakeholder-focused approach³. Patton focuses on a specific type of stakeholder, the primary intended user, which he defines as:

“...those specific stakeholders selected to work with the evaluator throughout the evaluation to focus the evaluation, participate in making design and methods decisions, and interpret results to assure that the evaluation is useful, meaningful, relevant, and credible. Primary intended users represent key and diverse stakeholder constituencies and have responsibility for transmitting evaluation findings to those constituencies for use.” (2008, p.72)

The literature includes insights on how to encourage effective stakeholder involvement. Patton emphasizes the need for the primary intended users to be very clearly defined at the beginning of an evaluation, and engaged throughout the evaluation process. This is, according to him, one of the most effective ways of ensuring evaluation use as, “intended users are more likely to use evaluations if they understand and feel ownership of the evaluation process and findings...” (2002, p. 1).

³ Several other authors support and demonstrate the importance of stakeholder involvement, including Alkhalaf (2007), Cullen et al. (2011), Laubli Loud and Mayne (2014), Balthasar (2009) and Marra (2000).

Alkin and Christie remind practitioners to think about both the depth and breadth of stakeholder involvement, emphasizing that both are necessary for successful use (2005, p. 119). Laubli Loud and Mayne propose rebranding evaluation as a tool for knowledge exchange more than knowledge generation, thereby accentuating the participatory focus of evaluation (2014, p.87).

Many authors have shared experiences of facilitating stakeholder involvement. For example, Marra (2000, p. 33) shares:

“The best ways to encourage the use of evaluation findings have been to involve the programme staff in defining the study and helping to interpret results, and to produce regular reports for the programme staff whilst the study is in progress. As Weiss (1998b) comments: ‘this kind of sustained interactivity transforms one-way reporting into mutual learning’ (p.30).”

Carden summarizes the point eloquently:

“We asked the users! And then we asked them again, and again, and stayed close to them throughout the study. But we did not only question. As we asked, we also gave. The value of the study emerged in these exchanges.” (2009, p.196)⁴

The case studies collected for this research, and available in Annex 4, relate similar experiences within the UN system.

Relevant Case Studies

- 3: GEF - Country-level evaluations
- 9: UN Women - Regional Mechanisms to Protect the Human Rights of Women and Girls
- 12: UNEP - Formative evaluation of the UNEP program of work
- 13: UNEP - Midterm evaluation of the Project for Ecosystem Services
- 17: UNICEF - National Child Protection Agenda in Thailand
- 18: UNICEF - Global Education Cluster
- 20: UNICEF - Progress evaluation of the Education in Emergencies and Post-Crisis Transition Programme
- 26: UNRWA - Steering Committee
 - 28: UNRWA – Evaluation of the agency’s Medium Term Strategy
- 32: WIPO – Knowledge-sharing evaluation

⁴ Knowledge to Policy

Discussion points

1. What can we learn from agencies that systematically engage and consult stakeholders? Are there lessons around which stakeholder engagement practices are most appropriate to various environments and evaluation types?
2. What can we do to improve the involvement and consultation of users and stakeholders in evaluations across the UN system?
3. Are there potential disadvantages to increased stakeholder participation in evaluation that may undermine the delivery of a good quality evaluation report?
4. What makes stakeholder engagement in the evaluation process meaningful (as compared to pro-forma)?

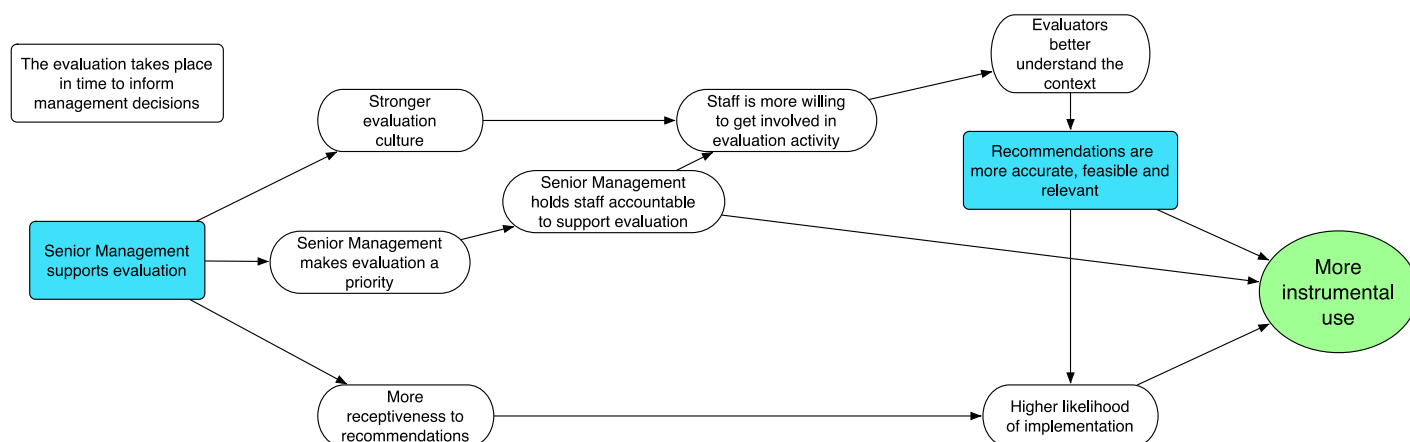
Key Message 2

The support of senior decision-makers is key, as is their commitment to the implementation of the recommendations.

In all organizations, support from the top helps drive change. Within most UN agencies, evaluation offices are not well positioned to make executive decisions. As such the support of senior decision-makers in championing evaluation activities is important to strengthen the evaluation culture; and make management more receptive to evaluation recommendations.

Mechanisms at play

Figure 2: Mechanisms linking senior decision-maker commitment and evaluation use



I. Senior management support helps strengthen the evaluation culture.

Beyond the championing or endorsement of evaluation activities, senior decision-makers (and in particular senior managers) have an important role to play to foster and support a strong evaluation culture in their organization. Without support from the top, it is unlikely that a strong evaluation culture will emerge. This culture is important in determining the demand for evaluations, and the extent to which stakeholders – in particular programme staff - engage in the process. Without such a culture, staff may be less willing to engage in the evaluation process, reducing participation in the evaluation and, as a result, weakening the quality of the recommendations generated.

II. By making evaluation activities a priority, senior managers encourage their staff to get more involved.

Evaluations which are timed to support planning or decision-making for a high-priority activity are more likely to be supported by senior managers. This support includes making their staff prioritize engagement with the evaluation, and holding them accountable for devoting sufficient time and energy to the evaluation activities. This pressure from above increases the likelihood of staff getting involved in the evaluation activities and this, as described in key message 1, greatly supports evaluation use.

III. When senior managers support evaluation, they tend to be more receptive to evaluation recommendations.

Positive attitudes towards evaluation and evaluators, and good understanding of the evaluation process and its potential benefits, makes senior managers more inclined to trust and embrace evaluation recommendations. This ultimately increases the likelihood that those recommendations are adopted.

Data

Respondents to our survey reported senior management support as the most important factor in ensuring that evaluation recommendations are implemented. 40% of respondents - and 50% of practitioners - listed “management leadership buy-in” as one of the main factors for evaluations to be used in decision-making processes.

About 80% of respondents reported the evaluation culture to be either an ‘important’ or a ‘very important’ factor for use. In open comments, they stressed how senior management plays a key role in fostering a strong evaluation culture, and how their “visible commitment, buy-in and proactive support are critical to ensuring that priority is given to effective evaluation”.

However, the message from interviews and open-ended responses to the survey was that, in practice, senior management is not always well-informed about evaluation and, as a result, does not foster sufficient support or demand. One interviewee did not consider “that senior management really takes the hard recommendations to the table”, while another claimed that “too often the most senior executive of any given UN entity is unwilling to listen and take action on the basis of these evaluations”. Overall, they reported low use among senior decision-makers.

Many interviewees reported working in an organization with a very positive, enabling evaluation culture. Others noted weaknesses in their organization's evaluation culture, in which project managers still feel threatened by evaluations, and recommended "a friendlier and less punitive strategy to promote evaluation should be used". A few interviewees considered that some high-level users are confused about the distinction between audit and evaluation, which exacerbates a 'fear' of evaluation and creates reluctance to use it. Another interviewee suggested that such high-level users need to be 'educated' about evaluation in order that they become more supportive of evaluation activities, and make good use of evaluation findings.

Notes from the literature

The literature broadly supports our key message. Patton explains that the most important factor for use is the 'personal factor', which he described as "... the presence of an identifiable individual or group of people who personally care about the evaluation and the findings it generates" (2008, p.66).

For the UN context, it seems that it is senior management who need to personally care about the evaluation, and that their support makes a very large difference to evaluation use. The support of senior decision-makers is also important in other contexts; for example, in a 2011 study of several foundations, Thompson and Patrizi found that, when an evaluation was led by the CEO, a lot more attention was paid to the findings.

The literature on use frequently mentions evaluation culture as an important factor in use, as it influences "the degree of support for evaluation" (Hawkins in Laubli Loud and Mayne, 2014, p.48). A strong evaluation culture serves several purposes: it helps to ensure that staff and stakeholders engage in the evaluation process; it supports the development of a rigorous approach to integrated performance measurement within an organization; and it increases the willingness of an organization to engage in evaluative activities, performance assessments and learning-oriented activities more generally. While we are not able to assess the strength of evaluation cultures across UN agencies, it does appear that the organizational context differs quite substantially across agencies and that this might influence evaluation use.

Characteristics of a strong evaluation culture:

According to Laubli Loud and Mayne 2014, an organization with a strong evaluation culture:

- (1) Engages in self-reflection and self-examination;
- (2) Engages in evidence-based learning;
- (3) Encourages experimentation and change; and
- (4) Encourages public reporting on its performance.

Finally, parts of the literature stress the importance of an evaluation policy (Hergueta, Schur and Thapa in Laubli Loud and Mayne, 2014, pp. 175-195). Interestingly, this isn't something that emerges from our study. Only one interviewee mentioned evaluation policy, and in the survey, the presence of an evaluation policy was ranked lowest in terms of suggested use factors.

Relevant case studies

- 1: ESCAP - Trust Fund for Tsunami
- 5: ICAO- Evaluation of Results-based management
- 14: UNESCO - Evaluation of UNESCO's standard-setting work of the culture sector
- 19: UNICEF - Independent review of UNICEF's operational response to the January 2010 earthquake in Haiti
- 26: UNRWA – Steering Committee
- 28: UNRWA – Evaluation of the agency's medium-term strategy
- 29: WFP - Transition from food aid to food assistance
- 30: WFP – Food assistance in Bangladesh

Discussion points

- 1. What are effective strategies to increase senior decision-maker support for evaluation activities?
- 2. In cases where there is no support, what can be done to ensure effective use?

Key Message 3

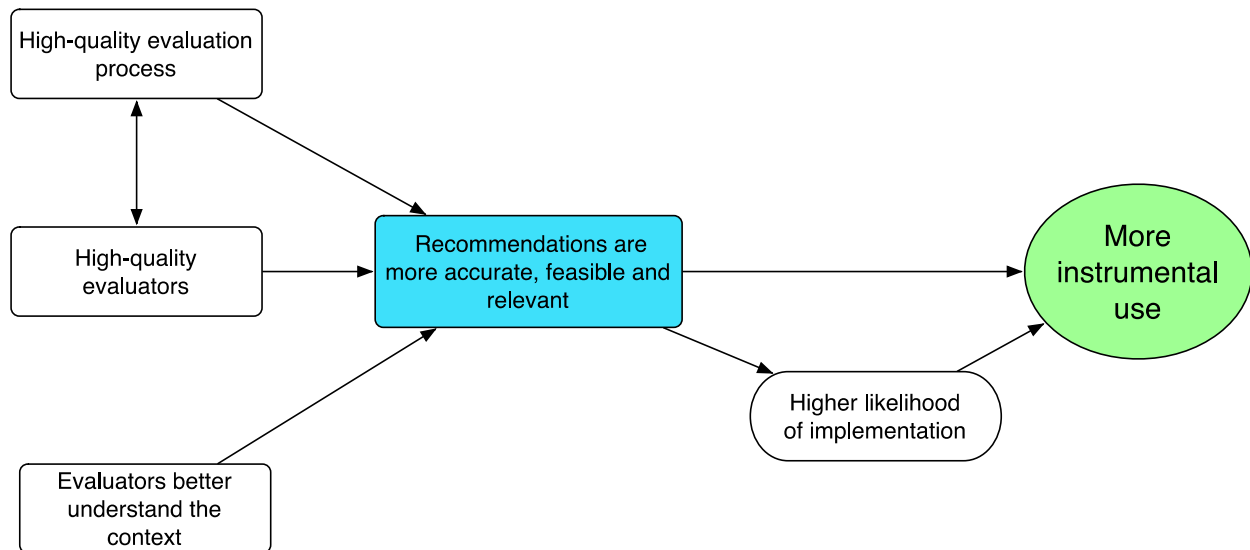
Evaluators need to ensure that recommendations are feasible and relevant.

Leaving aside process use, if an evaluation is to be useful, the recommendations generated have to be relevant and feasible. Indeed, the higher the quality of recommendations, the more likely that they will be accepted and implemented, and the use of the evaluation will increase.

Developing high quality recommendations is not simple, and depends on several factors including: user/ stakeholder involvement; a high-quality evaluation methodology; and high-quality evaluators.

Mechanisms at play

Figure 3: Mechanisms linking relevant recommendations and evaluation use



I. The quality of the evaluation process influences the quality of the recommendations.

Good recommendations can only emerge from a quality evaluation process. The quality of the evaluation process has several elements, and there is no standardized definition. We

define a ‘quality evaluation process’ as one with an appropriate methodology which produces timely recommendations.

In relation to methodology, we stress that the appropriate methodology should always be defined in relation to the context, the expertise of the evaluators, and the evaluation questions to be answered. While rigour is always important, there is not a single ‘gold-standard method’. Both quantitative and qualitative analytical tools can be adopted to suit a specific evaluation activity.

In relation to timeliness, in part this refers to the completion of the evaluation activities on time. More importantly, however, timeliness is about ensuring that evaluation activities are aligned with the programming cycle so that recommendations can feed into decision-making for the future of the programme.

II. The quality of the evaluators influences the quality of the recommendations.

Some argue that the quality of the evaluation process is closely related to the quality of the evaluators. We agree, but chose to discuss the quality of the evaluators as a separate component. As discussed above, good recommendations can only emerge from a quality evaluation process. This requires good evaluators. Evaluators that ask the wrong questions, do not successfully engage stakeholders, or infer the wrong conclusions from data would be unlikely to manage a quality process and come to quality recommendations.

III. Evaluators need to understand the context for the evaluation.

To produce feasible recommendations, the evaluators need to understand the realities of the commissioning organization, and internalize the various constraints – time, budget and otherwise – that could influence the uptake of their recommendations. Such insights typically come from regular engagement with the intended users and stakeholders, mentioned in key message 1 as one of the most important success factors for use. Involving stakeholders is a necessary step for evaluators to better understand the evaluation context, come to valid conclusions and make useful recommendations.

Data

From the survey, we learned that the quality of an evaluation was considered the most important factor supporting its use. All survey respondents rated the ‘quality of the evaluation’ as either ‘very important’ or ‘important’. While ‘quality’ remains a vague concept,

we believe that the it is very much a function of the quality of the evaluators and of the process, and in turn produces high-quality recommendations. Survey respondents and interviewees believed that the quality of the recommendations makes a very large difference to use, and there is a consensus around the importance of sufficient time for fine-tuning recommendations.

Once again, the reality of the UN context does not seem to reflect this ideal. One respondent reflects many similar comments when stating that, “the evaluators are sometimes disconnected from the realities of the organization and make recommendations that are not implementable”. In addition, some respondents considered that some evaluators were not sufficiently skilled, and that this ends up damaging the evaluation process. Interviewees suggested that the UN should improve efforts to recruit internal and external evaluators, and have a more thorough performance review at the point of the inception report.

Notes from the literature

Perhaps because this message appears self-evident to most, this topic is not much discussed in the literature. Hawkins confirms that “the quality of the end products of an evaluation is largely dependent on a good quality process” (in Laubli Loud and Mayne 2014, p.43). De Laat and Williams stress the importance of timeliness: “Without exception, the timeliness of evaluations and evaluation planning in relation to the programme or policy cycle turned out to be the most crucial factor for all types of evaluation use” (ibid, p.158). This also comes through our data, not as a very strong message, but rather as one of the ‘obvious’ factors that is generally agreed upon and therefore not considered to deserve further attention.

Relevant case studies

- 1: ESCAP - Trust Fund for Tsunami
- 6: ILO – Better Factories in Cambodia
- 10: UN WOMEN – Kenya evaluation of the gender and governance program
- 12: UNEP - Formative evaluation of the UNEP’s programme of work
- 17: UNICEF - National Child Protection Agenda in Thailand
- 25: UNRWA – Background paper
- 33: WIPO: Recommendations from IOD evaluation reports

Discussion points

1. What is the relationship between the various components of 'evaluation quality' (evaluators, process and timeliness)? Are they all factors necessary for the production of good recommendations in all cases, or are there special situations where not all are required?

Key Message 4

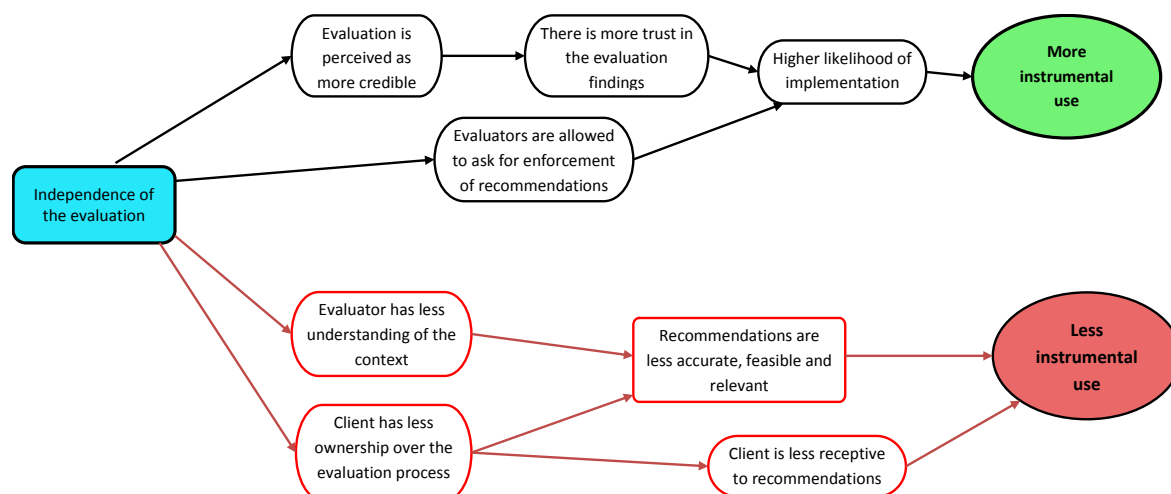
4. Independent evaluations must attempt to capture organizational realities.

Most professional evaluators agree that independence is a fundamental principle of evaluation, and it is increasingly seen as a key criterion of evaluation quality. It is a core principle of the OECD Development Assistance Committee (DAC) Quality Standards for Development Evaluation, and several international organizations – including the UN - have set up independent evaluation offices.

Independence enhances the impartiality and credibility of evaluations, thereby increasing the trust in, and support for, the resulting recommendations. However, there is a risk that the more independent the evaluator, the less they understand the context they are evaluating, potentially making recommendations less feasible and relevant. These tensions are explored in greater detail below.

Mechanisms at play

Figure 4: Mechanisms linking evaluation independence and use



I. Independence of the evaluation – or of evaluation functions – enhances the impartiality and credibility of evaluation activities.

The rationale behind having an independent evaluation function is clear: independence lends impartiality and credibility to the evaluation process and results. The theory is that users of evaluations, particularly senior decision-makers, might trust the findings of an

independent evaluation process more and, as a result, be more likely to implement the recommendations. Conversely, a perceived lack of independence can be used to dismiss evaluation findings on the grounds that they are not sufficiently credible.

Similarly, the independence of the evaluation function enables evaluators to ask senior management to enforce recommendations, increasing the likelihood that they will be implemented. Setting up an independent evaluation unit can also act as a signal to the wider organization that evaluation is important.

II. Independence could reduce evaluators' understanding of the context for the evaluation, and staff ownership over the evaluation process.

Independence can also have drawbacks, considering that truly independent evaluations are also external. First, external evaluators do not have as much organizational or programme knowledge as internal evaluators, thereby reducing their ability to understand the activity or programme being evaluated. Externally generated recommendations are often said to be less well-informed, relevant and feasible, as evaluators do not always understand internal dynamics and resource constraints. Given that relevant and feasible recommendations are considered a backbone of useful evaluations, it would therefore follow that independent evaluations have less use potential.

Second, programme staff and senior decision-makers typically have less ownership over independent evaluations, as they are not as involved in the commissioning process. As a result, they are more likely to consider independent evaluations as an external imposition. They may therefore show less willing to contribute to the evaluation process, or interest in the results, thereby reducing the potential quality and use of the recommendations.

Data

Interestingly, there appeared to be overall agreement on the role of independence within individual agencies, but not across agencies. This seems to indicate that an agency's evaluation culture influences practitioners' and users' opinions on this matter.

Most survey respondents agreed that the independence of evaluators is significantly important – over 90% rated independence as an 'important' or 'very important' factor for use. Notably, over 80% of intermediate users (senior management, donors, governing body) rated independence as 'very important', as compared to only 30% of evaluation practitioners, suggesting that senior decision-makers believe more strongly in the independence of evaluation. An interviewee echoed this finding, stating that “many

evaluators claim (independence is) not that important, but many stakeholders see it as important”.

Overall, interviewees had mixed responses on this topic. One disagreed with the sanctity of independence, explaining that “the evaluators are sometimes disconnected from the realities of the organization and make recommendations that are not implementable.” On the other hand, one intermediate user claimed that he “could not imagine how non-independent evaluations could ever be useful”.

While several respondents did see a trade-off between independence and the usefulness of recommendations, many others did not agree. Most consider that, while evaluations require a certain level of freedom of expression, it is absolutely essential for evaluators to understand the context very well. It appears that any trade-off that might exist has not fully been internalized, or recognized. Finally, one survey respondent noted in the open-ended questions that independence may be more important for accountability-driven evaluations than for learning.

Notes from the literature

Overall, the literature supports our finding that this is a complex topic with no clear answers. Most research recognizes that “the institutional arrangements for the planning and implementation of evaluations have a decisive influence on their utilization” (Balthasar 2009).

For the New Zealand context, Hawkins (in Laubli Loud and Mayne, 2014, p. 38) defines independence as: “the ability of evaluators to ‘speak truth to power’ without being compromised by the competing demands of policy and operation managers oriented toward meeting the demands of their executive team members, who in turn are focused on directives from their political leaders”. She also recognizes that “sustaining full independence without becoming too remote from policy and practice realities – as well as remaining relevant and useful – is a significant challenge” (Ibid, p. 41). According to her, there is a trade-off between making feasible recommendations and remaining independent.

De Laat proposes a useful framework to better understand the interplay between independence and usefulness. He points out that external evaluations can easily lose their independence, and that internal evaluators might actually provide less biased findings (in Laubli Loud and Maybe, 2014, p. 19).

Relevant case studies

- 19: UNICEF - Independent review of UNICEF's operational response to the January 2010 earthquake in Haiti
- 25: UNRWA – Background paper

Discussion points

1. What effective strategies have UN Independent Evaluation Offices adopted to ensure that independent evaluators do not become too remote from the reality of those they are evaluating?
2. What effective strategies have evaluation units without an independent status adopted to boost their impartiality and credibility?

Key Message 5

Management responses and follow-up processes must take place and be adequately supported.

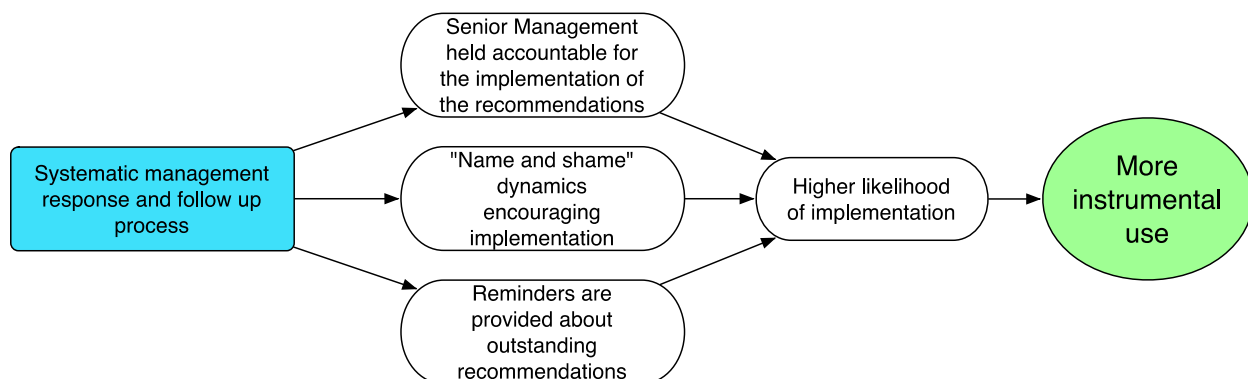
Systematic management responses and follow up processes increase the likelihood of implementation. The most important mechanism for the implementation of evaluation recommendations in the UN system is the management response, along with the associated follow-up process. Many UN organizations appear to have standardized, systematic ways to produce management responses, although this is not yet universal.

Standard 1.4 of the UNEG Norms and Standards for Evaluation in the UN System states that, to ensure that evaluation recommendations are utilized and implemented, “UN organizations should ensure appropriate evaluation follow-up mechanisms and have an explicit disclosure policy”. Standard 3.17 states, “Evaluation requires an explicit response by the governing authorities and management addressed by its recommendations”.

A 2010 document titled “UNEG Good Practice Guidelines for Follow up to Evaluations” suggests some elements of good practice for management responses. These include: (1) increasing the level of ownership of the evaluation findings prior to the management response; (2) clearly defining roles and responsibilities; (3) agreeing a deadline for the response; (4) nominating a focal point to coordinate the management response; (5) providing support by showing good examples of management responses; and (6) ensuring it is clear whether management accepts or rejects the recommendations.

Mechanisms at play

Figure 5: Mechanisms linking systematic management response and evaluation use



Processes for management responses and follow up ensure that there is a systematic way to ensure that the organization takes on board and deals with the recommendations of an evaluation. By involving senior management and publicising the findings, these processes can nudge the organization into implementing the recommendations, and prevents uncomfortable findings from being swept under the carpet.

I. Management responses and follow up processes keep senior management accountable for the implementation of evaluation recommendations.

Follow-up processes require periodic review of progress towards implementation, and for senior management to account for this progress – or lack thereof. This gives them greater incentive to take action towards implementation.

II. Management responses and follow up processes create a ‘name and shame’ dynamic, encouraging implementation.

Follow-up processes make progress towards the implementation of recommendations highly visible, creating a ‘name and shame’ dynamic whereby senior managers who fail to implement accepted recommendations can be clearly identified. To avoid that, managers are motivated to take action and implement recommendations.

III. Follow up processes remind managers about outstanding recommendations.

Busy managers might not be able to keep track of all the recommendations they or their predecessors have committed to implement. For that reason, follow-up processes are useful in reminding them of outstanding recommendations, thereby increasing the likelihood of those recommendations being implemented.

Data

Over 80% of survey respondents agreed that the follow-up to an evaluation is one of the key factors determining whether an evaluation is used. However, only 50% agreed that “there is a reliable follow up process to the evaluation recommendations after an evaluation is finalized”. A higher proportion – about 70% - agree that “management responses are reviewed by senior management/ governing bodies”, and a little under half report that “evaluation reports are discussed in-depth by senior decision-makers”. In many agencies follow-up is still not mandatory for all evaluations, and the follow-up process is not consistently applied. One interviewee suggested that responsibilities for follow-up needed to be more clearly defined.

All interviewees agreed that management responses are very useful in promoting use, but reported varying practices for implementing management responses. Some talked of serious and systematic consideration of recommendations, while others reported “no evidence that the evaluation has been taken seriously by senior management”. Some warned against management responses being taken too lightly, as “management response without follow-up and strong buy-in does not ensure that agreed recommendations get implemented”.

Finally, many interviewees reported difficulties in tracking the implementation of recommendations. For example, one board member interviewed said he’d like to know more about the extent to which past evaluation recommendations had been implemented, but didn’t know how to find this information. Indeed, practices for tracking the adoption and implementation of recommendations seem to vary widely. A best practice case in this regard is the Global Environment Facility (GEF), which maintains a Management Action Record, and submits an annual overview to the Council in its Annual Performance Report.

Notes from the literature

As management responses are UN-specific, there is very little on this topic in the wider literature. However, as noted above, UNEG has produced ‘Good Practice Guidelines for Follow up to Evaluations’ which, while very UN-centric, supports our key message.

Relevant case studies

- 16: UNFPA - Joint evaluation of the UNFPA-UNICEF joint program on Female Genital Mutilation
- 27: UNRWA – Interactive recommendation follow-up (27)

Discussion points

1. What effective strategies have been adopted to strengthen management response and follow-up processes? Have those been associated with more use?

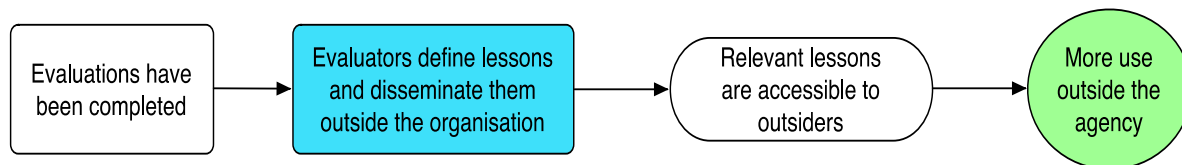
Key Message 6

Sharing of findings enables cross-organizational learning and use.

Generally, lessons learnt as part of evaluation activities tend to stay locked within an organization, as part of their knowledge management systems or their staff capacity. Freeing up lessons and findings from the confines of the individual organization would create the opportunity for evaluations to be used in other contexts and agencies.

Mechanisms at play

Figure 6: Mechanisms promoting cross-organizational evaluation use



- I. Evaluation use can be expanded if relevant lessons and findings are accessible to others than the primary intended users.

The more that is done to reach all potential users – which may go far beyond the immediate stakeholders – the more use can be generated from an evaluation. The kind of use generated through wider dissemination is more likely to be for learning than for instrumental use. Quite simple efforts to better target dissemination and ensure that evaluation lessons are accessible to external stakeholders can make a big difference to cross-organizational learning.

Evaluation practitioners should be encouraged to make a conscious effort to develop findings and lessons that might be relevant and applicable beyond the immediate evaluation stakeholders. Those lessons could potentially strengthen learning across teams, themes and organizations.

Data

Many interviewees felt that by more actively disseminating lessons from evaluations, evaluation offices could enable cross-programme, cross-thematic and even cross-organizational learning. However, in general respondents reported difficulties in both

accessing evaluation reports, and in particular in accessing the lessons learned from evaluations. Many survey respondents did not know where to find evaluation reports, even though most agencies have a central repository (often public) for evaluations.

However, there appears to be no easy way to access 'lessons learned'; one of the more useful and transferable components of evaluations. One survey respondent claimed that wider dissemination does not currently take place because "nobody feels responsible for extracting lessons", though there are some examples of innovative practice. The UNDP knowledge management team is planning to create a 'lessons learnt' database to enable wider dissemination of useful evaluation lessons. Similarly, UNEP attempted to better define and share lessons by preparing a 'Framework of Lessons from Evaluation'⁵. One interviewee suggested that there should be a requirement for new projects to take stock of previous evaluation findings on the topic

Notes from the literature

The literature does not focus specifically on cross-organizational learning, but more generally on improving dissemination. Myers claims that "dissemination, just like evaluation, is something that needs to be planned. A strategy that identifies the audiences for the sharing of evaluation findings, considers appropriate methods to achieve this, and develops appropriate timelines for it to be achieved will ensure a successful conclusion to this part of the evaluation journey" (2004, p.22). Sklar supports an active dissemination approach, which uses meetings supplemented by shortened reports as its main tools (2010, p.v).

Relevant Case Studies

- 8: IOM – Evaluation of Gender Mainstreaming
 - 28: UNRWA – Evaluation of agency's Medium Term Strategy

Discussion points

1. What type of tool or platform could better support cross-organizational learning? What type of information should be shared? What would be the best format for lesson sharing? What would be the relationship between lessons from evaluation and monitoring?

⁵ <http://www.unep.org/eou/Portals/52/Reports/Lessons%20Learned%20rpt.pdf>

2. In the context of information overload and thin evaluation capacity, would evaluators, programme staff and senior decision-makers actually have time to consult such a tool? In other words, would the tool be used?

Next Steps & Future Research

The messages we defined are for the most part simple and uncontroversial. Nevertheless, our data showed that the practices implicit in those messages are still far from being universally applied. We therefore believe it is worth re-emphasizing the importance of those messages and spreading the word, with the intention of improving evaluation use in the UN system. We encourage readers of this document to pilot some of the mechanisms that they feel are transferable to their situation and report back on their experience. The messages are useful in confirming some priority areas which could be explored to improve evaluation use in the UN system, and for further research.

Discussion points

In the previous sections, we defined several points for discussion. We use this section to summarize them and take stock of what our future research agenda might look like.

Key Message 1 – Points for Discussion

- What can we learn from agencies that systematically engage and consult stakeholders? Are there lessons around which stakeholder engagement practices are most appropriate to various environments and evaluation types?
- What can we do to improve the involvement and consultation of users and stakeholders in evaluations across the UN system?
- Are there potential disadvantages to increased stakeholder participation in evaluation that may undermine the delivery of a good quality evaluation report?
- What makes stakeholder engagement in the evaluation process meaningful (as compared to pro-forma)?

Key Message 2 – Points for Discussion

- What are effective strategies to increase senior decision-maker support for evaluation activities?
- In cases where there is no support, what can be done to ensure effective use?

Key Message 3 – Points for Discussion

- What is the relationship between the various components of ‘evaluation quality’ (evaluators, process, timeliness)? Are they all necessary for the production of good recommendations in all cases?

Key Message 4 – Points for Discussion

- What effective strategies have UN Independent Evaluation Offices adopted to ensure that independent evaluators do not become too remote from the reality of those they are evaluating?
- What effective strategies have evaluation units without an independent status adopted to boost their impartiality and credibility?

Key Message 5 – Points for Discussion

- What effective strategies have been adopted to strengthen management response and follow-up processes? Have those been associated with more use?

Key Message 6 – Points for Discussion

- What type of tool or platform could better support cross-organizational learning? What type of information should be shared? What would be the best format for lesson sharing? What would be the relationship between lessons from evaluation and monitoring?
- In the context of information overload and thin evaluation capacity, would evaluators, programme staff and senior decision-makers actually have time to consult such a tool? In other words, would the tool be used?

Research questions:

All of these discussion points can be summarised in the following three research questions:

- I. What are the specific mechanisms that make our key messages important? Are they context-specific? Do they always hold true?**
- II. What effective strategies have various evaluation offices or UN agencies used to strengthen their practice around each of our six key messages?**
- III. How can organizations be encouraged to take action to strengthen some of the practices and processes that support evaluation use?**

The S02 working group plans to centre our 2015-16 work plan around those main questions, with a clear view of improving evaluation use in UN agencies. More concretely, our plan includes the following activities:

- I. Validating our key messages with different audiences;
- II. Identifying synergies and coordinating with other UNEG working groups;
- III. Identifying gaps for good practice and prioritizing future research needs;
- IV. Conducting further research on the key messages that most need it;
- V. Producing a 'Good Practices for Evaluation Use' document, including findings from new and existing research;**
- VI. Holding consultations about the 'Good Practices for Evaluation Use' document and, subsequently, disseminating this document across UN agencies.

Bibliography

- Alie, R., Seita, J. (1997). Who's using evaluation and how? New study gives insight. *Nonprofit World*, Vol. 15, n° 5.
- Alkhalaf, A. (2007). The relationship among process use, findings use, and stakeholders involvement in evaluation.
- Alkin, M. C., & Christie, C. A. (2005). Unraveling theorists' evaluation reality. *New Directions for Evaluation*, 2005 (106), 111–128. doi:10.1002/ev.155
- ALNAP (2006), Review of Humanitarian Action Evaluation Utilisation. Development Institute, London.
- Antonie, R. (2011). The role of programme evaluation in the decision-making process.
- Balthasar, A. (2006). The effects of the institutional design on the utilization of evaluation: evidenced using Qualitative Comparative Analysis (QCA). *The International Journal of Theory, Research and Practice*, Vol. 12(3). http://www.interface-politikstudien.ch/downloads/deutsch/Art_Evaluation_QCA_e.pdf
- Balthasar, A. (2009). Institutional design and utilization of evaluation: a contribution to a theory of evaluation influence based on Swiss experience. *Evaluation Review* June 1, 33: 226-256. <http://www.ncbi.nlm.nih.gov/pubmed/18854493>
- Bamberger, M., Mackey, K. and Ooi, E. (2005). Influential Evaluations: Detailed Case Studies. Washington, DC: World Bank.
- Bates, R. (2004). A critical analysis of evaluation practice: the Kirkpatrick model and the principle of beneficence, *Evaluation and Program Planning*, Vol. 27: 341-347.
- Brousseau, A. and Contandriopoulos, D. (sd). *Evaluation Theories and Evaluation Use*. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3605074/pdf/nihms2438.pdf>
- Carden, F. (2009). Knowledge to policy. Making the most of development research: 16. Changing Practice: A Note on Using Evaluation for Organizational Change, Sage/IDRC, http://web.idrc.ca/en/ev-137284-201-1-DO_TOPIC.html
- Christie, C.A. (2006). Ensuring Evaluation Use. <http://www.eval.org/summerinstitute/06SIHandouts/SI06.Christie.BO23.Online.pdf>
- Church, C. and Rogers, M. (2006). "Chapter 10: Evaluation Utilization", in *Designing for Results: Integrating Monitoring and Evaluation in Conflict Transformation Programs*, <http://www.sfcg.org/Documents/dmechapter10.pdf>

- Cousins, J.B. (2003). Utilization effects of participatory evaluation. In T. Kelleghan, & D. L. Stufflebeam (Eds.), *International Handbook of Educational Evaluation* (pp. 245-265). Great Britain: Kluwer Academic Publishers.
http://link.springer.com/chapter/10.1007%2F978-94-010-0309-4_16
- Cracknell, B. (2001). The role of aid-evaluation feedback as an input into the learning organization. *Evaluation*, Vol. 7, n° 1: 132-145. http://www.stes-apes.med.ulg.ac.be/Documents_electroniques/EVA/EVA-GEN/ELE%20EVA-GEN%207536.pdf
- Cullen, A. et al (2011). The politics and consequences of including stakeholders in international development evaluation, *American Journal of Evaluation*, Vol. 32(3): 345-361.
- Delarue, J. et al (2009). Are evaluations useful? A review of the literature on “knowledge and decision-making”
- Education First Consulting (2010). Building the scale. Best practices for using evaluation to improve programmes and increase impact. <http://www.wcstonefnd.org/wp-content/uploads/2012/05/buildingtoscale.pdf>
- Ferry, M. and Olejniczak, K. (2008). The use of evaluation in the management of EU programmes in Poland.
- Grasso, P. (2003). What makes an evaluation useful? Reflexions from experience in large organizations. *American Journal of Evaluation*, Vol. 24: 507.
- Johnson, K. and Greenesid, L. (2009). Research on evaluation use: a review of the empirical literature from 1986 to 2005. *American Journal of Evaluation*, Vol. 30: 337-410.
- Hansson, F. (2006) Organizational use of evaluations. Governance and control in research evaluation. *Evaluation*, Vol. 12 (2): 159-178.
- Henry, G.T. (2000). Why not use? In V.J. Caracelli & H. Preskill (Eds.). *The expanding scope of evaluation use. New Directions for Evaluation*, No. 88, 85-98, San Francisco: Jossey-Bass.
- Henry, G. and Mark, M. (2002). A theory of evaluation influence: the multiple mechanisms and pathways through which evaluation can influence attitudes and action.
- Henry G. and Mark, M. (2003). Beyond use: Understanding evaluation’s influence on attitudes and actions. *American Journal of Evaluation*, 24,3: 293-314
- Herbert, J. L. (2014). Researching evaluation influence: A review of the literature. *Evaluation Review*, 38(5), 388–419. doi:10.1177/0193841X14547230

- Johnson, K. and Greenseid, L. (2009). Research on evaluation use: a review of the empirical literature from 1986 to 2005. *American Journal of Evaluation*, Vol. 30: 337-410.
- Kirkhart, K.E. (2000). "Reconceptualizing Evaluation Use: An Integrated Theory of Influence." In *The Expanding Scope of Evaluation Use*, ed. V.J. Caracelli and H. Preskill, 5-24. *New Directions for Evaluation* No.88. San Francisco: Jossey-Bass.
- Läubli Loud, M. and Mayne, J. (Eds.). (2014). *Enhancing evaluation use: insights from internal evaluation units*. Los Angeles: SAGE Publications, Inc.
- Levin-Rozalis, M. and Rosenstein, B. (2005). The changing role of the evaluator in the process of organizational learning. *Canadian Journal of Program Evaluation*, Vol. 20, n^o 1: 81-104.
- Leviton, L.C. (2003). Evaluation use: advances, challenges and applications. *American Journal of Evaluation* December 1, 24: 525-535.
- Marra, M. (2000). How much does evaluation matter? *Evaluation*, Vol. 6 (1): 22-36.
http://www.stes-apes.med.ulg.ac.be/Documents_electroniques/EVA/EVA-GEN/ELE%20EVA-GEN%207460.pdf
- Myers, P. and Barnes, J. (2004). *Sharing evaluation findings: disseminating the evidence*.
- Olejniczak, K. (2009). What is evaluation and how it can be measured?
http://www.esparama.lt/ES_Paramama/angliskas_medis/programming_for_2007_2013_tee/about_the_programming/files/Evaluation_Conference_2009/2.1.0_Olejniczak_Evaluation_utilization_260309.pdf
- Patton, Michael Quinn. (2008). *Utilization-Focused Evaluation* (4th Edition). SAGE Publications.
- Picciotto, R. (2005). "Use of Evaluation Findings to Improve Development Effectiveness: Panel Discussion." In *Evaluating Development Effectiveness*, ed. G.K. Pitman, O.N. Feinstein, and G.K. Ingram, 347-54. World Bank Series on Evaluation and Development, vol.7. Piscataway, NJ: Transaction Publishers.
- Rist, R. and Morra Imas, L. (2009). *The Road to Results: Designing and Conducting Effective Development Evaluations*. Washington, DC: World Bank, Independent Evaluation Group.
- Rangasami, J. and Richter, T. (2011). Case study of a recommendations-based evaluation focused on utility.
<http://www.impactconsulting.co.za/downloads/SAMEA%20presentation%20Recommendations-based%20study.pdf>

- Ruskus, J. and Alkin, M. (1984). Factors common to high-utilization evaluations.
- Sklar, M. (2010). The influence of dissemination techniques on different evaluation stakeholders. http://sdsu-dspace.calstate.edu/bitstream/handle/10211.10/336/Sklar_Marisa.pdf?sequence=1
- Smith, N. and Brandon, P. (2008). Fundamental Issues in Evaluation. New York. Guilford Press.
http://uqu.edu.sa/files2/tiny_mce/plugins/filemanager/files/4281330/bird/Fundamental%20Issues%20in%20Evaluation.pdf
- Stockbauer, U. What makes evaluations useful?
http://ec.europa.eu/regional_policy/sources/docconf/edimbourg/pdf/stock_en.pdf
- Stewart, F. (2005). "Evaluating Evaluation in a World of Multiple Goals, Interests, and Models." In Evaluating Development Effectiveness, ed. G.K. Pitman, O.N. Feinstein, and G.K. Ingram, 3-28. World Bank Series on Evaluation and Development, vol.7. Piscataway, NJ: Transaction Publishers.
- Thompson, E. and Patrizi, P. (2011). Necessary and not sufficient: the state of evaluation use in foundations. <http://evaluationroundtable.org/documents/Necessary-and-Not-Sufficient-Final.pdf>
- United Nations Evaluation Group. (2005). Standards for Evaluation in the UN System. Retrieved from <http://www.uneval.org/document/detail/21>
- United Nations Evaluation Group. (2010). UNEG Good Practice Guidelines for Follow-up to Evaluations. Retrieved from <http://www.uneval.org/document/detail/610>
- Weiss, C. (1977). Using Social Research in Public Policy Making. Laxington, MA: Lexington Books.
- Weiss, C. and Bucuvalas, M. (1980). Truth tests and utility tests: Decision-makers' frames of reference for social science research. *American Sociological Review*, 45, 302-313.
- Weiss, C. (1988). "Evaluation for Decisions: Is Anybody There? Does Anybody Care?" Pp. 171-84 in Debates on Evaluation, edited by Marvin Alkin. Newbury Park, CA: Sage. Reprinted American Evaluation Association keynote address originally published in *Evaluation Practice* 9(1):5-19.
<http://www.sciencedirect.com/science/article/pii/S0886163388800175>
- Weiss, C. (1998). Have we learned anything new about the use of evaluation? In: *American Journal of Evaluation* 19(1), 1998 pp.21-33.

Weiss, C. (2004). Identifying the intended uses of an evaluation. International Program for Development Evaluation Training (IPDET) presentation, Ottawa, July.
http://www.irdc.ca/ev_en.php?ID=58213_201&ID2=DO_TOPIC.

World Bank, Operation Evaluation Department. (2005). Influential Evaluations: Detailed Case Studies. Available at www.worldbank.org/ieg/ed.

Xavier, R., Mehta, A. and Gregory, A. (2006). Evaluation in use: the practitioner view of effective evaluation.
http://www.prismjournal.org/fileadmin/Praxis/Files/Journal_Files/Evaluation_Issue/XAVIER_ET_AL_ARTICLE.pdf

Annex 1. Research Methodology

Objectives

The objective of this research project was to better understand how UN agencies use evaluation and the factors that support and hinder evaluation use. More specific objectives for this project were to:

1. review the literature related to evaluation use and assess the extent to which conclusions and recommendations could apply to the UN setting;
2. interrogate UN evaluation users and practitioners to capture patterns and perceptions of evaluation use in the UN setting;
3. ask the wider international evaluation community about their experiences and perceptions of use;
4. capture UN case studies about positive instances of evaluation use;
5. share recommendations and findings across UN agencies; and
6. define a programme of work for S02 going forward.

Data

Most of the data-gathering tools were finalized before the research consultant came on board. An attempt was made ex-post to ensure that the data gathered as part of the exercise was supportive of the conceptual framework for the project, also designed ex-post.

Literature Review

The review covered the most relevant pieces of academic and organizational writing on evaluation use. Over 50 academic sources (journal articles, books and doctoral theses) were reviewed and summarized by the consultant. This information fed into the definition of the conceptual framework for this project, as well as some sections of the main document.

Online Survey

An online survey was designed and implemented using the Survey Monkey platform. Over 140 current or past UN employees and consultants with relevant experience of evaluation – as users or practitioners – were invited to respond to the survey. Those individuals were nominated directly by evaluation experts from the UN agencies collaborating in this project.

In addition to the UN survey, a similar survey was posted on the forum M&E News, inviting external evaluators to answer questions about evaluation use.

Semi-structured interviews

A questionnaire was developed prior to the hiring of the research consultant and used as the basis for semi-structured interviews with various UN staff involved in evaluations in different capacities. The questionnaire covered most topics included in the project's conceptual framework.

Some of the interviews were conducted by the UNEG research consultant and others by UNEG counterparts in the respective agencies.

Case Studies

Our UNEG counterparts were asked to prepare case studies reflecting instances when evaluations were used in their agency. In January 2015, the SO2 Working Group provided additional guidance for these case studies, asking for them to be shaped around the following three sections:

1. What were the objectives of the evaluation and who were the intended primary users?
2. How was the evaluation used?
3. Why was the evaluation used and successful?

Those case studies were analyzed and lightly edited for this report. We selected the strongest case studies of two types:

1. case studies showcasing the key messages identified in this report; and
2. case studies showcasing a successful example of evaluation use.

The analysis and triangulation of these four data sources formed the basis of the key messages defined in this report.

Research Questions

The main research questions were defined, refined, reworded and further focused in the development of our data collection instruments. The questions, asked as part of the online surveys, interviews and case studies, are the following:

1. In your organization, is evaluation used?
2. In your organization, how is evaluation used?
3. In your organization, who uses evaluation?
4. What are the factors in place supporting/hindering use?
5. What are practices that have been successful at generating more use?

Survey instrument

Annex 5. QUESTIONNAIRE (next page) QUESTIONS ON:	Ask to*			1 Practitioner 2 Immediate user 3 Intermediate user
1. PROCESS	Involvement, capacity building, quality of process; relationship & trust building between evaluator and users			
	1	2	3	
1.1.1 When do you consider commissioning evaluations?	x			
1.1.2 When do you consider using evaluations? Why?		x	x	
1.2 Are your evaluations designed primarily for learning or accountability? (Express a %)	x			
1.3 What are the types of use of evaluations that are carried out by your agency: use for...? (i) Strategic decision-making (ii) Summative / programmatic / institutional (iii) Formative: programme / project design and review (iv) Use external to organization (for national partners, for development partners, or professional groups) (v) Use to enhance added value (e.g. via meta-analyses and lessons notes), contribute to evaluation profession and advance evaluation methodology and approaches	X X X X X			
1.4 What are incentives to motivate potential users to demand and/or use evaluations as input for their own decision-making processes?	X			
1.5 Does independence make a difference in the use of the evaluation, or other aspects such as quality...?	x	x	x	
1.6 Does the involvement of stakeholders (e.g. through reference groups etc.) in the evaluation process affect the use of evaluation recommendations? If yes, how?	x			
2. UTILITY	Utility of evaluation and the recommendations (relevance for decision making, timely etc.)			
2.1 Which stakeholders are most interested in demanding / which in using evaluations? Why? 2.1.1 Which stakeholders should demand the evaluations, but don't? Why?	X x	X	x	
2.2 Are reports and results easily accessible and effectively / transparently disseminated? 2.2.1 Are reports publicly available?		X	X x	
2.3 Are evaluations being demanded?	X	X	X	
2.4 Are you aware of evaluations that may be used as an input for decision making? If yes, how are you made aware?	x	x	x	
2.5 To what extent are evaluation reports considered by senior management / boards? 2.5.1 Are responses to the recommendations considered or reviewed by senior management / executive boards?	x	X x	X x	
2.6 What do you believe are the factors (mention a few, i.e. 3-5) that make an evaluation most useful?	X	X	X	

2.6.1 What are the top 2 items to be improved to make evaluations used more often?	X	x	x	
2.6.2 How do you measure the utility of evaluations?	x			
2.7 Which are incentives for potential users to demand and/or use evaluations as input for their decision-making?	x	x	x	
2.8 What should evaluators do during planning, conduct, follow up to evaluations to make it more useful for you?		x	x	
2.9 When did the knowledge transfer take place, mainly during the conduct of the evaluation, at the closing stages or through follow up on the recommendations?	x	x	x	
3. COMMUNICATION	During planning, process, results, recommendations channels and tools, presentation of preliminary finding			
3.1.1 Are you getting stakeholders on board, making them aware of the usefulness of the evaluation early on? How?	x			
3.1.2 Are you made aware of the upcoming evaluation of potential use for you? If yes, how?		x	x	
3.2 Are the evaluation entity, planning office, and finance office communicating well to ensure that evaluations are considered in next round of planning cycle? If yes, how?	x			
3.3 Is the evaluation function active in creating demand for use of evaluations? If yes, how? If not, why not?	x	x	x	
3.4 Within your context is the public asking information from evaluation commissioners? Explain who, and why.	x			
3.5 Do donors demand evaluations for their decision making? Do they actually use evaluations for it?	x	x	x	
4. FOLLOW UP	Follow up: process of follow up, formal or informal			
4.1 What happens after an evaluation is conducted?	X	X	X	
4.1.1 Is there follow up after an evaluation is finalized?	X	X	X	
4.1.2 If so, what are the steps of this follow up?	X			
4.1.3 Who monitors and leads the follow up?	X			
4.1.4 For how long follow up happens after an evaluation?	x	x	x	
5. ACTUAL USE	Actual use of evaluation / impact			
5.1 Who are primary (intended/unintended) users of evaluations in the context of your agency?	x			Answer and Fill in Annex 2
5.2 Did use of evaluations by users produce real changes related to the evaluation purposes?	x			
5.3 How relevant/pertinent you (user) find the evaluation?		x	x	
5.4 Does using evaluations produce changes in policy actions?	x	x	x	
5.5 What do you believe are the 3 major factors that make an evaluation useful? (E.g. relevance of findings, timing of the evaluation, etc.)	x	x	x	

5.6 What were the characteristics of those you did use?	x		x	
5.7 How were these reflected in your most successful two evaluations?	x			
5.8 Why were other evaluations less useful?	x		x	
5.9 How would you consider the actual use of evaluations by yourself? (a) High (systematic use for decision making, new project cycles, corrective actions); (b) Medium (sometimes used, ...); (c) Low (ad-hoc use, not so frequent); (d) Not used at all. 5.9.1. To what extent are they being used for accountability (%), learning (%) or other purposes (% - explain which other purposes)?	X x	X x	X x	
5.10 How would you consider the actual use of evaluations by your organization / agency / institution? (a) High (systematic use for decision making, new project cycles, corrective actions); (b) Medium (sometimes used, ...) (c) Low (ad-hoc use, not so frequent); (d) Not use at all.	x	x		

Annex 2. Use of Evaluations – Theory and Data from the UN System

1. Levels of use

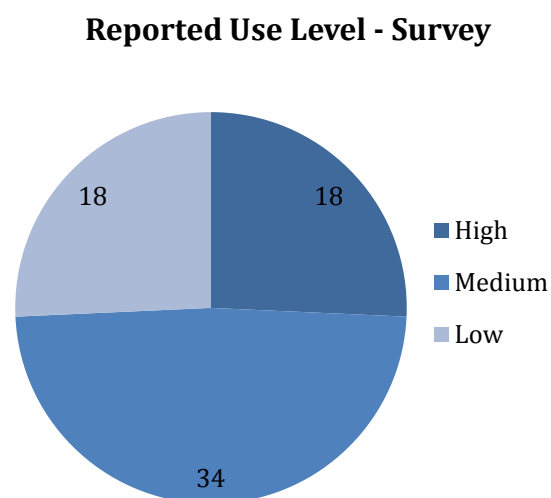
There has been a worldwide surge in demand for evaluation (Patton, 2008, p. 29). All evaluations are done with the intention that they are used. Anything else would be absurd. Indeed, Patton (1997, in Herbert 2014, p.389) claims that evaluations that are not used, regardless of their quality, tend to be considered failures. The primacy of ‘use’ as a criterion to judge the success of evaluations has been debated (Henry, 2000; Weiss, 1988). For the purposes of this document, it is not necessary to agree whether use is the single most important judgment criterion. but simply that evaluation use is important, and that more should be done to enable it.

In reality, a large share of evaluations are not used, or are not used enough (Laubli Loud and Mayne, 2014; Patton 2008). In 1988, Carol Weiss stated that “even the best and greatest evaluations only minimally affect how decisions get made”. While much might have changed since then, more recent literature suggests that there is still a long way to go for evaluations to be as useful as they could, or should be.

From our data:

The primary data collected as part of the S02 exercise support the findings of evaluation researchers: like them, most of our respondents agree that it is important for evaluation to be used. However, as described in the literature, the reported level of actual use is variable, both within and across organizations. As can be seen in the chart below, a little over a quarter of the survey respondents rated use as ‘high’ (which we defined as ‘systematic use for decision making, new project cycles, corrective actions’) or ‘low’ (which we defined as ‘infrequent use’). About half rate use as medium (which we defined as ‘periodic use’).

Different respondent groups (we asked practitioners, immediate evaluation users and intermediate evaluation users) responded differently. ‘Medium’ was consistently the most frequent response, but 40% of immediate users reported use to be high, while only 12% of intermediate users made the same claim. Whether



those reported levels of use point to differences in perception, or to actual different levels of use across organizations, is hard to tell.

Nevertheless, we can conclude that there is a lack of systematic evaluation use, which is what we would ideally have observed. Many of the respondents we spoke to shared anecdotal evidence of use, but did not think that use was very high overall. Many thought that evaluation was most often used when it supported actions that were already planned by decision-makers – when change was already intended. There was a general perception that more could be done to generate greater and better evaluation use.

On a positive note, no interviewee or survey pointed to any instance of evaluation misuse, and most respondents agreed that their agency, or the UN in general, were on an ‘upward use trajectory’, getting better and better at ensuring that evaluations are used.

The patterns observed support Henry and Mark, according to whom “knowledge utilization in decision-making is not automatic”, which implies that “a deliberate and committed strategy is required from organizations in order to move beyond sporadic successes” (2003, p. 298). This supports the conclusions of this exercise, which provided some anecdotal evidence of successful use by certain users, but found that others did not know whether evaluations had really been used, or strongly believed that more could be done to generate better use. Overall, in the UN system, evaluation use does not yet appear to be systematic.

2. Defining evaluation use

The literature reviewed, by and large, explores the roots of this under-use, as well as options to rectify the problem. Before we discuss the factors that support or hinder use, we need to introduce a few definitions of concepts around evaluation use that support our discussion. We start by adopting this definition of evaluation use:

“Evaluation use, or evaluation utilization, occurs when evaluation information in the form of findings, or evaluation practice, has influence on the actions or thoughts of stakeholders.” (Alkin, 2005, p. 143 in Alkhalaf 2007, p.8).”

Researchers disagree about the specific meanings of ‘use’ and ‘utilization’. Some “believe that the term ‘use’ implies direct use of evaluation findings”, in contrast to ‘utilization’, which refers to “a dynamic process that occurs over time” (Patton, 2008, p. 107 in Alkhalaf 2007). Others claim the opposite, saying that “utilization holds linguistic connotations related to direct and instrumental use only” (King, 1982, Weiss, 1980 in Alkhalaf, 2007). The jury is

still out on which term - 'use' or 'utilization' - is narrower in sense - focusing only on direct action like the implementation of recommendations - and which incorporates learning-oriented outcomes, such as knowledge creation. As a solution to this debate, Kirkhart suggested substituting the terms 'use' and 'utilization' with 'influence', a term he meant to better capture all types of use, both direct and indirect, action-oriented and learning-oriented (Kirkhart, 2000, p.7).

In this document, we choose to use the term 'use', which we intend to capture all types of evaluation influence discussed above.

3. Different purposes of evaluation

All evaluations aim to cover both accountability and learning. The literature often makes the distinction between 'accountability-driven' and 'learning-driven' evaluations – as two different *purposes* of evaluation (Laubli Loud and Mayne 2014, p. 6). According to most authors, whether an evaluation is driven by accountability or learning purposes influences the configuration of the evaluation, as well as the type of use to which it will be put.. Accountability is essential to ensure that development or humanitarian interventions can be implemented in the medium term, as accountability-driven evaluations provide evidence that projects or programmes are achieving results. However, the utility of an accountability system depends on who is held accountable, by whom, for what and how (Patton, 2011, p.119). Consequently, accountability-driven evaluations are described as being typically less conducive to use, whereas learning-driven evaluations are more likely to engender instrumental and conceptual use (de Laat in Laubli Loud and Mayne; Hawkins in Laubli Loud and Mayne 2014; Patton 2008).

For the purpose of this project, while we do discuss accountability-driven evaluations, we put a greater emphasis on learning-driven evaluations, as these are more typically associated with use. We acknowledge that the inclusion of both types of evaluation might result in lower levels of use being reported overall, but we did not consider it realistic to exclude accountability-driven evaluations from the data gathering.

From our data:

Survey responses clearly show that accountability is still the main driver of evaluations in most agencies. The most frequent response to the question “in your organization, what percentage of evaluations are done for accountability purposes, and what percentage for learning purposes?” was 60%/40%. In addition, over 60% of respondents agreed that the biggest factor in the commissioning of evaluations was that evidence on results or

performance was needed, again demonstrating the prominence of accountability-focused evaluations.

On the other hand, half of respondents also agreed that evaluations were commissioned when programmatic improvements needed to be identified, suggesting strong intended use for those evaluations. Indeed, while accountability-driven evaluations have the potential to be used for more than the direct demonstration of accountability, they are typically not designed for that, thereby reducing their use potential.⁶

4. Users and types of use

It is widely recognized that evaluation can have multiple types of use (Laubli Loud and Mayne, 2014, p.3). Comprehensive typologies of evaluation use feature several types of use⁷, including types of political use and misuse of evaluations. For this report, we chose to focus on three direct types of use:

- IV. Instrumental use;
- V. Conceptual use; and
- VI. Process use,

placing greater emphasis on instrumental use.

Instrumental Use

The most widely known use type is ‘instrumental use’, or most simply ‘direct use’. This type of evaluation use is the most commonly referenced, and the one most people think about when referring to evaluation use. In the UN system, it can be conceptualized as the adoption and implementation of an evaluation recommendation.

Conceptual Use

Conceptual use happens “when an evaluation influences how key people think about a programme or policy; they understand it better in some significant way, but no action or decision flows from the findings” (Patton 2008, p.103). It happens when evaluation findings provide a new way of thinking about an issue, new insights into a programme, but do not generate any specific action or change. In the UN system, this relates to the evolving conceptualization and understanding of those who design strategies and programmes.

⁶ While the survey gives information on different purposes of evaluation (learning vs. accountability), it doesn’t distinguish between the different kind of use..

⁷ We recommend consulting Patton’s typology of use, (2008, pp. 112-113), which is very comprehensive.

Process Use

Process use “refers to cognitive, behavioural, programme and organizational changes resulting, either directly or indirectly, from engagement in the evaluation process and learning to think evaluatively” (Patton 2008, p. 108). It occurs when stakeholders “learn from the evaluation process itself or make programme changes based on the evaluation process rather than findings” (Ibid, p. 109). In the UN system this mainly relates to changes to implementation or programming resulting from interactions between the evaluation team and key stakeholders.

Different evaluation users

The literature identifies several types of users. For the purpose of this exercise, we chose to focus on two types of users most commonly encountered in the UN system: people with decision authority over a programme, and those with direct responsibility for it.

2. **People who have decision authority over the programme** include policy makers, funders, and advisory boards (*Definition based on Greene, 2006 in Patton, 2008, p. 61*). In the UN context, this user group is composed of Senior Management, Governing Bodies, Donors and Partner Governments. It is responsible for approving evaluation recommendations, and is ultimately responsible for their implementation.

People who have direct responsibility for the programme include programme developers and managers, administrators in the organization implementing the programme, and direct service staff (*Definition based on Greene, 2006 in Patton, 2008, p. 61*). In the UN context, this user group is composed of programme staff. This group is responsible for supporting the evaluation process and taking action to implement the recommendations approved by senior decision-makers.

According to Patton (2008) and Carden (2009), every evaluation should very clearly define its ‘primary intended users’. These are defined as “those specific stakeholders selected to work with the evaluator throughout the evaluation to focus the evaluation, participate in making design and methods decisions, and interpret results to assure that the evaluation is useful, meaningful, relevant, and credible. Primary intended users represent key and diverse stakeholder constituencies and have responsibility for transmitting evaluation findings to those constituencies for use” (Patton 2008, p. 72). In the UN system, the primary intended users of evaluation are most often individuals from the second group mentioned above, with direct responsibility for and knowledge of the programmes being evaluated.

From our data:

In the UN system, many evaluations are initiated by donors. More than half of our survey respondents identified donor request as one of the main factors behind the commissioning of evaluations.

That said, the most frequently cited primary users of evaluation were programme staff (identified by about 50% of respondents), followed by management and senior management (about 20% each). Interestingly, the respondents who listed senior management as primary users were among the most likely to rate evaluation use in their organization as either 'high' or 'low'.

The most frequently cited reasons for use were for 'programme improvements' and 'strategic decision-making'. This is positive, as primary users largely appear to be stakeholders with the power to act upon findings and recommendations. Those who report using evaluation results for operational, programmatic, project planning and design are about 50% less likely to report low use from evaluation.

5. Evaluation in the UN system

Within the UN system, evaluation is conducted at three levels: (a) system-wide or secretariat-wide; (b) by the central evaluation offices of UN organizations; and (c) by decentralized evaluation functions within some UN organizations. It should be noted that the decentralized evaluation functions are much closer to management structures, and evaluations designed at this level are more likely to be designed for learning purposes.

6. Use factors

A myriad of factors facilitating evaluation use are discussed in the literature – in fact we found over 100 factors across more than 20 sources⁸. We attempted to simplify the literature by grouping those 100 factors together and, as a result, produced the following summary of use factors. We isolated eight individual use factors, which we grouped into two categories: factors related to the context of the evaluation, and factors related to the evaluation activities. These eight factors formed the basis of the conceptual framework used for the data collection exercise undertaken in this project. Those interested in the full conceptual framework can find more detail in Annex 3.

⁸ Full bibliography available in the Bibliography section

Table 2: Summary of use factors

Context in which the evaluation takes place	
	Evaluation culture
	Organizational structure and incentives
	Characteristics specific to the evaluation activity
Evaluation activities	
	Evaluation design and process
	User/stakeholder involvement
	Evaluator qualities
	Evaluation product
	Post-evaluation process

While the table above demonstrates well the range of factors that can affect evaluation use, we found that our respondents did not tend to give all of those factors equal relevance. For example, user/ stakeholder involvement was very strongly stressed by our respondents, whereas the area of organizational structure and incentives were barely mentioned. Respondents focused more on the specific issue of independence and its impact on the evaluation process and product.

We originally intended to use the above framework as the basis for our main discussion, but decided to change course and instead focus on some of the messages that emerged very strongly from the data we collected. Some of these messages are very closely related to factors on the above list, but others are not. The following section describes the key messages we received from our respondents.

Annex 3 Detailed Conceptual Framework

Factors (potentially) associated to use and demand		
Context in which the evaluation takes place		
	Evaluation Culture	
		<i>Support and commitment by senior management/Strong evaluation champions</i>
		<i>Presence of evaluation advisory groups</i>
		<i>Good organizational understanding of evaluation</i>
		<i>Presence of evaluation policy</i>
		<i>Respect for/visibility of the evaluation function</i>
		<i>Quality of the M&E infrastructure in place</i>
		<i>Evaluative thinking/use of evaluation information in decision-making</i>
		<i>Previous positive experience with evaluation/demonstrated use</i>
		<i>Availability and awareness of evaluation reports and products</i>
	Organizational Structure and Incentives	
		<i>Strength and position of the Evaluation Unit</i>
		<i>External pressures for accountability/information</i>
		<i>Root of the demand for evaluations</i>
		<i>Systems in place to feed evaluation findings into decision-making processes</i>
		<i>Incentives to learn</i>
		<i>Incentives to act on evaluation recommendations</i>
		<i>Ability of the organization to implement recommendations</i>
		<i>Organizational stability/Human resources</i>
		<i>Resource availability</i>
	Characteristics Specific to the Evaluation Activity	
		<i>Presence of evaluation champion for the specific evaluation</i>
		<i>Nature of the subject being evaluated</i>
		<i>Resources dedicated to this specific evaluation</i>
		<i>Presence of individuals/networks/intermediaries/brokers to facilitate the evaluation</i>
		<i>Timing of the evaluation in the policy cycle</i>
		<i>Political climate</i>
		<i>Main users' capacity/ability to receive and implement findings</i>
		<i>Nature of the evaluation findings</i>
Evaluation Activities		
	Evaluation Design and Process	
		<i>Involvement of potential users at an early stage</i>
		<i>Planning for timeliness in the policy cycle</i>
		<i>Planning communication and dissemination early in the evaluation process</i>
		<i>Appropriateness and relevance of the evaluation approach</i>
		<i>Rigor and quality of the design, data sources and triangulation</i>

		<i>Encouragement of stakeholders to reflect critically on the project</i>
	Stakeholder/User Involvement	
		<i>User involvement from the beginning</i>
		<i>Mechanism(s) for user involvement at all stages of evaluation</i>
		<i>User involvement in defining evaluation questions</i>
		<i>Participatory relationship between evaluation staff and programme staff</i>
		<i>Sharing preliminary findings with users</i>
		<i>Face-to-face meetings with producers and users of knowledge take place</i>
		<i>Direct participation of users in evaluation teams</i>
	Evaluator Qualities	
		<i>Personal qualities: independence, credibility, openness, flexibility, adaptability</i>
		<i>Procedural qualities: communication and facilitation skills</i>
		<i>Technical competences</i>
	Evaluation Product	
		<i>Timeliness</i>
		<i>Relevance of findings</i>
		<i>Quality/validity of findings</i>
		<i>Report is user-friendly/ final deliverables are tailored to different audiences</i>
	Post-Evaluation Process	
		<i>Dissemination/communication strategy is sound and implemented</i>
		<i>Formal system for managers to respond to findings</i>
		<i>Formal follow-up system to verify whether recommendations have been implemented</i>
		<i>Evaluators maintain significant involvement in follow-up activities</i>
		<i>Findings are actively disseminated to users</i>
		<i>Meta-evaluation and/or interagency lesson learning activities take place</i>

Annex 4 - Case Studies on Successful Use of Evaluation

In this section, we present some case studies that were submitted as part of this exercise. In the table below, we list all case studies submitted that illustrated successful evaluation use. When applicable, we list the key message supported by the case study. Readers interested in the case studies can find them all in the following pages, on the page specified in the table.

Case Study Number	Agency	Case Study	Key Messages (if applicable)
1	ESCAP	Trust Fund for Tsunami	Key Message 2, 3
2	FAO	Evaluation of FAO's role and work in statistics	
3	GEF	Country-level evaluations	Key Message 1
4	GEF	Fifth Overall Performance Study	
5	ICAO	Evaluation of results-based management	Key Message 2
6	ILO	Better Factories in Cambodia	Key Message 3
7	ILO	Sustaining Competitive and Responsible Enterprises	
8	IOM	Evaluation of gender mainstreaming	Key Message 3, 6
9	UN WOMEN	Regional Mechanisms to Protect the Human Rights of Women and Girls	Key Message 1
10	UN WOMEN	Kenya evaluation of the Gender and Governance Programme	Key Message 3
11	UNDP	Assessment of Development Results in Uruguay	
12	UNEP	Formative evaluation of the UNEP program of work	Key Message 1, 3
13	UNEP	Midterm evaluation of the Project for Ecosystem Services	Key Message 1
14	UNESCO	Evaluation of UNESCO's standard-setting work of the culture sector	Key Message 2
15	UNESCO	Evaluation/ review of the UNESCO Education Category I Institutes	
16	UNFPA	Joint evaluation of the UNFPA-UNICEF Joint Programme on Female Genital Mutilation	Key Message 5
17	UNICEF	National Child Protection Agenda in Thailand	Key Message 1, 3

18	UNICEF	Global Education Cluster	Key Message 1
19	UNICEF	Independent review of the UNICEF operational response to the January 2010 earthquake in Haiti	Key Message 2, 4
20	UNICEF	Progress evaluation of the Education in Emergencies and Post-Crisis Transition Programme	Key Message 1
21	UNICEF	Joint evaluation of the UNFPA-UNICEF Joint Programme on Female Genital Mutilation	
22	UNICEF	Evaluation of the UNICEF emergency preparedness systems	
23	UNICEF	Transforming Residential Institutions for Children and Developing Sustainable Alternatives	
24	UNICEF	Civil Registration Support in Cameroon	
25	UNRWA	Background paper	Key Message 3, 4
26	UNRWA	Steering Committee	Key Message 1, 2
27	UNRWA	Interactive recommendation follow up	Key Message 5
28	UNRWA	Evaluation of agency medium-term strategy	Key Message 1, 2
29	WFP	Transition from Food Aid to Food Assistance	Key Message 2
30	WFP	Food Assistance in Bangladesh	Key Message 2
31	WFP	School Feeding in The Gambia	
32	WIPO	Knowledge sharing evaluation	Key Message 1
33	WIPO	Recommendations from IOD evaluation reports	Key Message 3

Case Study 1: ESCAP - Trust Fund for Tsunami

Full Title	ESCAP Trust Fund for Tsunami, Disaster and Climate Preparedness in the Indian Ocean and Southeast Asian Countries
What were the objectives of the evaluation and the intended primary evaluation users?	<p>The evaluation assessed the ESCAP-established Multi-donor Voluntary Trust Fund on Tsunami Early Warning Arrangements in the Indian Ocean and Southeast Asia. This fund had been established in late 2005 when the lack of a tsunami early warning system for the Indian Ocean was made evident through the Indian Ocean tsunami of December 2004. An Advisory Council makes policy and funding decisions for the Fund.</p> <p>The intended primary users of the evaluation were: firstly, the donors, comprising the Government of Thailand, as foundation donor, and the Governments of Sweden, Turkey, Bangladesh, and Nepal; and secondly, ESCAP.</p> <p>The purpose of the evaluation was threefold: to account for results to stakeholders of the Fund; to assess future scenarios for the Fund in terms of focus, role, funding and governance; and to generate useful recommendations related to policy issues and management of the Fund, including scenarios for possible future donations and governance.</p>
How was the evaluation used?	<p>Following the evaluation, the Advisory Council of the Fund was briefed on the recommendations. ESCAP formulated a management response with follow-up actions, which was endorsed by ESCAP's Executive Secretary.</p> <p>In response to the evaluation recommendations, ESCAP took several actions to establish a more focused strategic direction for the fund, and to increase cohesiveness and coherence across the various projects, donors, programme managers, and implementers. Among the actions taken are:</p> <ul style="list-style-type: none"> • meetings between selected key stakeholders are now convened to build trust and explore synergies; • a joint calendar is now compiled each quarter for all projects to support joint collaboration where possible; • a strategy summary was prepared by ESCAP and agreed by the Advisory Council; • a concrete resource mobilization and communication plan is implemented. <p>On the whole, the evaluation showed positive performance of the trust fund and was therefore used for advocacy purposes.</p>
Why was the evaluation used and successful?	<p>Several factors came together to contribute to a successful outcome.</p> <p>First, the quality of the evaluation was high. It thoroughly consulted the relevant stakeholders, which established the necessary trust and ownership. The recommendations were useful, and while detailed enough, manageable, lending themselves to implementation. The quality of the evaluation was due to the high level of M&E expertise of the project staff involved in managing the evaluation, the fact that the project staff prioritized the evaluation process</p>

	<p>among competing management tasks, and the profile and skills of the evaluator, who was a recognized authority in disaster risk reduction with significant evaluation as well as senior management experience.</p> <p>Second, the evaluation was demand-driven and welcomed. Stakeholders shared the assumption that the trust fund would need to change and evolve over time, and it had therefore been a management decision to regularly conduct evaluation of the trust fund. The evaluation was a welcome tool to manage the direction and process of the change. Through its strong summative focus, the evaluation succeeded in quantifying some of the benefit of the Indian Ocean Tsunami Early Warning System, and could point to concrete results for the Fund. This, along with the fact that the results not only pointed to areas for improvement, but also contained praise for the trust fund, was important in gaining enthusiasm, credibility and buy-in among donors and senior management. Acceptance of the evaluation was further facilitated by the evaluator's overall profile, and her standing among disaster risk reduction experts.</p> <p>Third, the conditions for implementing the changes were favourable, on the one hand because the group of staff working on the trust fund is small and could therefore flexibly adjust, and on the other hand because ESCAP's top management backed the process and helped ESCAP implement the changes.</p>
--	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Case Study 2: FAO - Evaluation of the FAO role and work in statistics

Full title	Evaluation of the FAO role and work in statistics
What were the objectives of the evaluation and the intended primary evaluation users?	FAO statistics are widely used to conduct work and decision-making both within and outside FAO. The evaluation was aimed at assessing the relevance, quality and utility of the FAO role and work in statistics from a users' perspective. It was commissioned by the FAO governing bodies and management, and was carried out in close consultation with programme staff and partner agencies.
How was the evaluation used?	Evidence gathered by the evaluation revealing major weaknesses in global and country level data and capacities were used as a basis for redesign. This included the establishment of new governance structures for coordinating the FAO statistical system, the development of the first-ever FAO statistical programme, the formulation of global and regional capacity development strategies, the development of a FAO Statistical Quality Assurance Framework and the launch of IT projects for the reengineering of FAOSTAT and the development of an FAO data warehouse.

Why was the evaluation used and successful?	Some of the factors for success included: (i) a strong interest by FAO governing bodies and management on the topic; and (ii) the quality and timing of the evaluation.
----------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Case Study 3: GEF – Country-Level Evaluations

Full title	Country-level evaluations
What were the objectives of the evaluation and the intended primary evaluation users?	Country-level evaluations cover GEF support across GEF Agencies, projects, and programmes in a given country or in a cluster of countries. These evaluations assess the performance and results of GEF support at country level, and how this support is linked to national environmental and sustainable development agendas, as well as to the GEF mandate of generating global environmental benefits within its focal areas. The target audiences for country-level evaluations are the GEF Council, national stakeholders including GEF Focal Points in the countries involved, GEF agencies, and the GEF Secretariat. Since 2008, results of country-level evaluations have been aggregated in Annual Country Portfolio Evaluation Reports (ACPERs) presented yearly to the Council. Since the introduction of country-level evaluations in 2006, the evaluation office has conducted 23 country-level evaluations across all the GEF geographic regions in the world. Through these evaluations and alongside thematic, performance and impact evaluations, the Office has helped to shape and contribute to new policies that define GEF today.
How was the evaluation used?	A main feature of country-level evaluations (CLEs) is the focus on issues that are important from the perspective of country stakeholders. For example, the Turkey and Moldova CLEs, summarized in the ACPER 2010, elevated the issue of involvement of GEF Operational Focal Points (OFPs) in project monitoring and evaluation (M&E). OFPs tended to be actively involved with GEF agencies until obtaining the OFP project endorsement letter, a requirement for submitting the project proposal to the GEF. OFPs were not involved further during implementation. M&E information did not always flow from GEF Agencies to national partners and vice versa, and the role of the national partners in M&E processes was limited. Based on recommendations of ACPER 2010, the Council requested GEF agencies to systematically involve OFPs in M&E. Subsequently a new minimum requirement was added in the review of the GEF M&E Policy, on engagement of OFPs in M&E plans, activities, mid-term reviews, and final evaluations. The quality-at-entry review of GEF projects presented in the Office's APR 2012 found that new projects started to specify how OFPs would be informed and involved in M&E activities.

	<p>Another example of a core issue for country stakeholders has been the limited resources available to support countries in tackling land degradation. The Cameroon, Egypt, and Syria CLEs, summarized in the ACPER 2009, concluded that there was a significant gap in resources available for combating land degradation in those countries. The ACPERs showed that countries didn't receive the resources and support in land degradation they were expecting, either directly through land degradation focal area allocations or as multifocal area projects. In parallel, the mid-term review of the Resource Allocation Framework (RAF) of the GEF discussed the need to introduce one integrated resource allocation system for all GEF focal areas per country. All these evaluations were instrumental to inform the Council discussions and decision-making at a time when the resource allocation system, the RAF, was to undergo a major review in preparation for GEF-5. Land degradation was the only GEF focal area with Global Environmental Benefits (GEBs) comparable across GEF member countries in all geographic regions, which made it possible to add it to biodiversity and climate change focal areas in the revised country resource allocation system, now called the System of Transparent Allocation of Resources (STAR). Later on, the Fifth Overall Performance Study of the GEF (OPS5) concluded that the land degradation focal area drew more resources than expected, exceeding its original allocation under GEF-5, which confirmed once again how crucial land degradation is to national stakeholders.</p>
<p>Why was the evaluation used and successful?</p>	<p>One of the main factors facilitating the influence of country-level evaluations is the inclusiveness of the process, especially when it comes to hearing the voices of country stakeholders. The Office approach to country-level evaluations includes the systematic engagement of GEF stakeholders throughout the whole evaluation. Engagement starts from discussing the scope of the terms of reference. It continues during the data gathering and analysis phase and culminates in discussing preliminary findings and concrete areas for improvement before the evaluation independently reaches firm conclusions and recommendations. This inclusive process allows country stakeholders to have ownership over the evaluation of their GEF portfolio. It also increases significantly the evaluation use in the country while at the same time – as we have seen above – recommendations to the GEF leads to decision-making by the GEF Council, which results in institutional change. The Office is currently exploring new modalities for further engaging stakeholders in its ongoing country-level evaluations (i.e. through online stakeholder consultation platforms and ad hoc webinars during the evaluation phase), as well as joint country-level evaluation modalities with country governments to further increase the use of those evaluations at country level.</p>

Case Study 4: GEF - Fifth Overall Performance Study

Full title	Fifth Overall Performance Study of GEF (OPS5)
What were the objectives of the evaluation and the intended primary evaluation users?	<p>Overall Performance Studies of GEF are independent external evaluations conducted every four years to inform the next replenishment cycle. These studies assess the extent to which GEF is achieving its objectives and identify potential improvements. OPS5 was conducted between March 2012 and November 2013 to inform the replenishment for GEF-6.</p> <p>The effort was led and implemented by the GEF Independent Evaluation Office (IEO) and involved the staff, junior and senior consultants, consulting firms and institutions. OPS5 also included several quality assurance mechanisms: a team of senior independent advisors representing developing and developed nations; and a reference group consisting of members of the evaluation offices of the GEF agencies. OPS5 included stakeholder interaction during Extended Constituency Workshops organized by the GEF Secretariat, and via targeted consultations, especially with the members of the GEF CSO network.</p> <p>OPS5 targeted mainly stakeholders with a governance role in the GEF (the replenishment group, the Council, and the Assembly), as well as stakeholders responsible for implementations of decisions made by the governing bodies (mostly, the Secretariat and the GEF Agencies).</p>
How was the evaluation used?	<p>As reflected in the Summary of Negotiations of the GEF Trust Fund's Sixth Replenishment (GEF/R.6/26), the OPS5 provided "an important context for the discussions" during the replenishment process. It also informed and already started contributing to policy changes that are influencing the work of GEF-6 (2014-2018). A few examples of such on-going changes include: current revisions of the results-based management system, approval of the gender action plan to implement the gender mainstreaming policy, changes in the co-financing policy.</p>
Why was the evaluation used and successful?	<p>There were a few factors that contributed to the success of OPS5:</p> <ol style="list-style-type: none"> 1) Timing: the study was planned to provide evaluative evidence during the replenishment process. Unlike previous overall performance studies, the IEO prepared and presented interim findings in the earlier stages of replenishment negotiations – the first report in April 2013, the progress report in September 2013 - the final conclusions and recommendations were also presented to the replenishment group (in December 2013), and to the Council and the Assembly (in June 2014). 2) Relevance: OPS5 provided information on the key issues of the GEF business and results model, as well as the key institutional policies. 3) Credibility of the GEF Independent Evaluation Office: two professional peer reviews of the evaluation function in the GEF (2009 and 2014) confirmed high credibility of the IEO in the GEF network as defined by the expertise, independence, and degree of transparency of the evaluation work. According

	<p>to the Second Professional Peer Review (2014) Overall Performance Studies are considered to be very useful by Council members and adequately meet needs in terms of accountability for the replenishment of the GEF fund.</p> <p>4) Communication: OPS5 findings were communicated with the use of various types of products: several reports – first, progress, and final report; 21 technical documents; video explaining the process behind and views of the evaluators of some of the conclusions; and infographics summarizing main conclusions and recommendations in a concise and clear manner.</p> <p>At the same time, timing was also a limiting factor: the replenishment participants requested that future overall performance studies be presented even earlier in the replenishment process, as the final conclusions and recommendations of OPS5 were presented close to the end of negotiations, which limited their use in decision-making.</p>
--	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Case Study 5: ICAO - Evaluation of Results-Based Management

Full title	Evaluation of results-based management at ICAO
What were the objectives of the evaluation and the intended primary evaluation users?	<p>The evaluation was intended to provide a review to the ICAO governing body and senior management on progress made in implementing results-based management at ICAO. The period covered by the evaluation was from 2004 to 2013, that is, the period since the issuance of a resolution of the ICAO Assembly requiring the Organization to introduce a more strategic, results-oriented approach to business planning, as a basis for developing the Organization's budget. The intended primary users of the evaluation were the ICAO Council and the Secretary General.</p>
How was the evaluation used?	<p>Following the completion of the evaluation, the recommendations were presented to and discussed by the ICAO Finance Committee and the Council. This led to an Organization-wide effort, facilitated by the Finance Branch, to formulate performance indicators for each major organizational unit, as the first step in creating a Corporate Performance Management System at ICAO. The Organization's planning and budget document and process have also been improved as a result of the evaluation.</p> <p><u>Formulation of performance indicators</u></p> <p>The organizational units were asked to dedicate time to formulate performance indicators and submit these to the Finance Branch, which coordinated the process. A user-friendly orientation guide on performance management using RBM was developed by the Finance Branch and disseminated to staff. A network of focal points for all organizational units was</p>

	<p>established, which facilitated the process within the respective units. In addition, workshops facilitated by the Strategic Planning Officer were held with each organizational unit to raise awareness on performance management, assist with the formulation of performance indicators and provide quality assurance. Although the initial stage in the formulation was to submit indicators at the output level, units were encouraged to internally measure and consider outcome indicators. A final list of key performance indicators (KPIs) is to be submitted to the ICAO Council in 2015, which will be invited to select those KPIs it may wish to monitor. It is expected that measurements of organizational performance that will be collected thereafter will feed into the planning of the subsequent budgetary cycle (i.e. the 2017-2019 triennium).</p> <p><u>Planning document and process improvements</u></p> <p>The planning and budget document will consolidate the performance indicators submitted. The process was facilitated by the accompanying workshops, the orientation guide on performance management using RBM and by the Secretary General's endorsement of the performance management framework. The framework is being rolled out progressively as a joint effort between the Secretariat and the Council, and it is acknowledged to be a resource intensive but worthwhile exercise.</p> <p>Whereas it is too premature to assess the impact of the evaluation, and some recommendations were not accepted, a number of positive and concrete steps were implemented immediately following the completion of the evaluation.</p> <p>The key recommendations of establishing a performance management framework and providing performance management training were accepted and expected progress has been made in these areas since the evaluation. For example, performance management workshops were held for Council and Secretariat staff, facilitated by the Director of Cabinet of the Secretary General and the Strategic Planning Officer of the Finance Branch.</p>
Why was the evaluation used and successful?	<p>The fact that the evaluation used the Joint Inspection Unit's (JIU) results-based management model, which ICAO endorsed in 2004, as the benchmarking framework ensured an accountability link with the JIU and the Secretariat to the Council, and strengthened the credibility of the evaluation.</p> <p>Council Members, among other stakeholders, appreciated being consulted as part of the evaluation. These are influential stakeholders who champion performance management in the Organization. In particular, two Council Members were considered to have championed performance management and helped bring attention to and raise awareness on the importance of improving performance management at ICAO. They have supported the evaluation and its results. In particular, they have advocated that ICAO establishes a performance management framework, which was the key recommendation of the evaluation, and was subsequently accepted by the ICAO Secretariat.</p>

	<p>The timing of this evaluation at ICAO was also appropriate for several reasons. With the arrival of an additional evaluation officer, expectations for evaluation output by the Evaluation and Audit Advisory Committee and the Council were heightened. Also, considering the zero-nominal growth of the Organization's budget and a mind-set of "doing more with less", the topic of performance management was of particular interest to the ICAO Council.</p>
--	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Case Study 6: ILO – Better Factories in Cambodia

Full title	Better Factories in Cambodia (BFC) – Midterm Cluster
What were the objectives of the evaluation and the intended primary evaluation users?	<p>The evaluation was intended to assess the progress of 11 projects initiated under the Better Factories Cambodia programme. The goal of the programme is to reduce poverty by expanding decent work opportunities in the garment export industry and contribute to the growth of exports through promoting socially responsible production and compliance with ILO core standards, as well as Cambodian labour law.</p> <p>Clients of the evaluation were the donors, the Better Factories Cambodia project management team, the ILO Country Office Director for Cambodia, ILO technical experts, as well as tripartite stakeholders in the employers and trade union organizations of the apparel industry in Cambodia.</p>
How was the evaluation used?	<p>Following the results of the evaluation, recommendations were presented to project management and the centralized HQ Better Work managers. The recommendations offered suggestions for correction of strategies and strengthening of those areas which were not making adequate progress. Substantial improvements in the direction of Better Factories in Cambodia were made on two fronts:</p> <p><u>Policy Strengthening</u></p> <p>The Better Work programme monitoring policy was seen by the evaluation to play out in too tough a manner. Management response to the recommendations resulted in BFC restructuring staff to accommodate more advisory services and assistance mechanisms that could focus on factory systems and root cause analysis. In this way, the monitoring policy will be less of a "rating" policy and become more active as an advice mechanism, to include a "satisfaction survey" pilot programme.</p> <p><u>Programme adjustments</u></p> <p>1) One of the key underlying principles of ILO is to ensure that there is fruitful collaboration between governments, workers and employers in all ILO projects. This cluster evaluation noted that tripartism was not working at optimal levels due to charges of bias. Management response noted that</p>

	<p>immediate changes in the current phase of the project would likely not be corrected sufficiently, but the programme document for the second phase was modified to strengthen the role of a national steering committee to facilitate better equality of cooperation and participation amongst ILO partners in the advisory committee and implementation of technical cooperation.</p> <p>2) A recommendation resulted in an expansion of the projects' engagement with buyers. Findings of the evaluation showed that there was potential to work more extensively with buyers, and management response resulted in a more innovative interaction with international buyers to expand collaboration on issues such as addressing the re-emergence of child labour and assisting employers in their own auditing activities.</p> <p>3) These multiple projects generated an enormous amount of information on training services, collaboration amongst partners and advocacy. One of the recommendations suggested that a more structured "public disclosure" strategy address the need for better presentation of these issues through a comprehensive website that could: i) assess progress on BFCs impact on the apparel industry; and ii) publish the results of its "factory satisfaction" surveys and case studies. Management response started these initiatives under a public disclosure strategy launched and sustained during the remaining phases of the projects.</p> <p>4) All ILO evaluations are meant to assess gender issues, and one recommendation of the evaluation resulted in new monitoring protocols for gender discrimination to be introduced into performance committees to correct perceived inequalities in training opportunities.</p>
Why was the evaluation used and successful?	<p>ILO is taking advantage, when possible, of clustering evaluations of similarly themed projects. This evaluation is a very good example of the way in which a cluster evaluation can make more efficient use of the evaluation effort, and also spread knowledge more effectively of what is working across a number of projects. In this way, evaluation findings can benefit a larger complement of activities, more efficiently focus strategic use of evaluation findings and engender a better overall technical understanding of what ILO is learning from its interventions.</p> <p>This evaluation also shows how a rigorous and timely management response can optimize the benefits of a midterm evaluation for improving project direction. There was a strengthening of collaboration to correct imbalances in participation amongst ILO's traditional stakeholders, as well as the introduction of a public disclosure strategy, which was completely new. This meant that staff responsibilities needed to be adjusted, that committee mandates needed to be realigned with core principles, and innovations for collaboration were initiated.</p>

Case Study 7: ILO - Sustaining Competitive and Responsible Enterprises

Full title	Sustaining competitive and responsible enterprises (SCORE I)
What were the objectives of the evaluation and the intended primary evaluation users?	<p>This project's performance was reviewed based on the following six evaluation criteria: (i) relevance and strategic fit of the intervention; (ii) validity of the intervention design; (iii) intervention progress and effectiveness; (iv) efficiency of resource use; (v) effectiveness of management arrangements; and (vi) impact orientation and sustainability of the intervention. It will also mainstream gender equality.</p> <p>Clients of the evaluation were the donors SECO and NORAD, the SCORE project management team, ILO Country Office Directors, ILO technical experts, as well as tripartite members of the Global Project Advisory Committee and National Committees and partner organizations in the evaluated countries.</p>
How was the evaluation used?	<p>ILO conducted a final workshop where the draft report of the evaluation was shared amongst the stakeholders of the evaluation, allowing them provide corrections and comments on the evaluation. The evaluator was given a list of consolidated comments before finalizing the evaluation.</p> <p>Following the completion of the evaluation, recommendations were presented to project management and the centralized HQ department Director who coordinated the global management response. The recommendations offered suggestions for improvements for the second phase of the project, which was to begin shortly after the evaluation was presented to management. Results included, inter alia, the following:</p> <p><u>Policy Strengthening</u></p> <p>SCORE was perceived by the evaluator to be a policy approach to enterprise development that should be "branded" by the ILO, and the evaluation recommended that the country experience gained should provide a "cross-fertilization" of ideas to improve ILO's branding and marketing efforts in this regard. This required adjustments to and improvements for SCORE's results-chain logic and its monitoring and evaluation capacity. Management response was able to introduce market assessments in some of the countries and a revised results-chain, including M&E training. Part of the knowledge gained through the cross-fertilization approach resulted in a recommendation for the next phase of SCORE to pursue a public-private partnership policy. This was adopted through management response and a formal agreement with one of the stakeholders was established.</p> <p><u>Programme adjustments</u></p>

	<p>1) As is often the case with complex projects, the evaluator found that activities and their implementation would benefit from longer project duration for a second phase, and this was positively accepted by management, who revised the second phase timing and duration.</p> <p>2) The evaluation suggested an adjustment to terminology, especially in the context of finalized a global branding effort. Management then developed a revised glossary which was shared on the project's website.</p> <p>3) The evaluator recommended that subsequent adjustments to knowledge sharing, training and workshops would need to take place. Citing a methodology from a related ILO project (WISE), which was branded by ILO in the 90s, the evaluator suggested improving the overall peer learning mechanism through involving small- and medium enterprise (SME) owners and senior managers. Work on this was begun as part of management response, with an experimental model introduced in Ghana.</p> <p>4) All ILO evaluations are meant to assess gender issues, and one recommendation resulted in a new draft gender strategy planned for adoption in the second phase of the project.</p>
Why was the evaluation used and successful?	<p>Large-scale global evaluations of multiple country programmes offer the opportunity for all countries involved to participate, share knowledge and gain insights from different national perspectives. A particular benefit of this evaluation was that the evaluator was able to utilize findings from previous related evaluations, to better frame the recommendations on intervention branding. As this was a high budget intervention, it had already undergone a mandatory midterm evaluation and the stakeholders in the project were able to benefit from that experience as they participated in the final assessment. Due to the level of understanding of the evaluation process, projects such as this tend to experience more timely turn around on the management response, especially because a second phase of the same project was in the pipeline, for which the evaluator could direct recommendations. The second phase of the project document was adequately revised – through the management response exercise – to reflect the conclusions and findings of the phase one final evaluation. Additionally, further innovations to the project concept were recommended in two specific project strategies that utilized findings from previous evaluations and incorporated good practice identified from the final phase of the projects.</p>

Case Study 8: IOM - Evaluation of Gender Mainstreaming

Full title	Evaluation of the gender mainstreaming policy and strategy in IOM
What were the objectives	The main objective of was to evaluate the overall performance and achievements of IOM in mainstreaming gender in its programme activities and in promoting gender balance in its staffing policy, including an assessment of

of the evaluation and the intended primary evaluation users?	the impact of the IOM gender policy on migration management. The evaluation identified good and bad practices and examined the extent to which the policy for both components was implemented either in isolation or institutionalized (looking at management responsibilities and organizational structure, allocated resources, application of formal rules and guidelines, capacity building activities, collaboration and partnership among others). The intended primary users were IOM as a whole, its Member States and donors.
How was the evaluation used?	This evaluation report has the peculiarity of having a 'second life'. Conducted in 2006, it evaluated the status of the IOM gender mainstreaming policy and strategy 10 years after it had been adopted in 1995. The recommendations were taken into account and the report was extensively discussed inside IOM as well as with donors and Member States. 10 years later, in 2015, it was used again as a benchmark for measuring progress on the implementation of recommendations, in the framework of a new review of IOM gender policy, and to compare the strengths and weaknesses identified today with those identified 10 years ago.
Why was the evaluation used and successful?	At the time of first publication, the evaluation met the interest of the audience listed above. It was the first time that IOM had embarked on a thematic and strategic evaluation related to its gender policy (and culture). It provided interesting evidence and original perspectives on how to reinforce gender, taking into account IOM specificity. The report and recommendations were found to be highly relevant, objective, sometimes innovative and very useful. This is why, 10 years later, much of its content appeared to be still globally relevant and the evaluation had a second life. This is more possible with strategic and thematic evaluations than with more traditional programme evaluations.

Case Study 9: UN WOMEN - Regional Mechanisms to Protect the Human Rights of Women and Girls

Full title	Evaluation of the project Regional Mechanisms to Protect the Human Rights of Women and Girls in South East Asia
What were the objectives of the evaluation and the intended primary	The UN Women Regional Office for Asia and the Pacific (ROAP) undertook an evaluation of the project, which involved two regional human rights bodies from the Association of South East Asian Nations, covering 10 countries in South East Asia. The project had a Steering Committee (comprising UN Women management and the donor) since its inception in 2010 that functioned as an evaluation management group for the final evaluation. Following the inception meeting of the evaluation process, UN Women ROAP set up a stakeholder reference group to oversee the evaluation planning and progress. However, because the key partners were

evaluation users?	intergovernmental bodies, it was difficult to limit the number of stakeholders in the reference group due to the politically sensitive nature of relationships. Therefore, a participatory approach was taken, and ROAP broadened the group to ensure comprehensive stakeholder participation to review the report findings and recommendations. All concerned stakeholders provided their feedback to the report and provided their inputs into the recommendations ranking them as high, medium and low priority.
How was the evaluation used?	These recommendations and the discussions from this meeting were inputs into the next phase of the programme.
Why was the evaluation used and successful?	<p>The participatory approach, engaging all stakeholders, facilitated buy-in, quality control, management of expectations, disclosure of findings, sharing of recommendations as well as ownership in implementing the recommendations.</p> <p>The evaluation process - where an independent group of people spoke to the stakeholders and then shared their findings with them, as well as the involvement of the stakeholders in designing the next phase of the project - was important in cementing the relationship of trust built by UN women with ASEAN, important in the context of the new UN Women- ASEAN Memorandum of Understanding.</p>

Case Study 10: UN WOMEN - Kenya Evaluation of the Gender and Governance Programme

Full title	Evaluation of the Gender and Governance Programme III (GGP III)
What were the objectives of the evaluation and the intended primary evaluation users?	GGP III represented the third phase (2009-2013) of the GGP Programme, launched in September 2004. The programme worked to promote equal opportunities and access to services for men and women, addressing the need to include women's issues in governance structures, and support women's leadership at national and local levels. The overall goal of GGP III was to ensure that Kenyan women and men are able to access services and opportunities and exercise their rights equally. UN Women partners included more than 40 civil society organizations and key government agencies like the Ministry of Gender, Children and Social Development; National Commission on Gender and Development; the Electoral Commission of Kenya and government institutions ⁹ .

⁹ Please note the Ministry of Gender, Children and Social Development is now listed as the Department of Gender under the Ministry of Devolution and Planning; the National Commission on Gender and Development has been disbanded and a National Gender and Equality Commission (NGEC) formed; the Electoral Commission of Kenya (ECK) is now the Independent Electoral and Boundaries Commission of Kenya (IEBC).

How was the evaluation used?	As this was the final phase of the GGP with no possibility of extension, UN Women decided to incorporate the evaluation recommendations into the next UN Women programming cycle. Additionally, as the evaluation coincided with the planning cycle for both the government of Kenya and the United Nations in Kenya, the evaluation findings and recommendations were availed for use as inputs to these processes. Once the draft evaluation report was issued, UN Women convened a broad stakeholders meeting (beyond the reference group) to discuss and validate the draft recommendations, and draft a roadmap for their implementation. UN Women then organized a consultative planning workshop for the UN Women Strategic Note 2014-2018. During this meeting, various stakeholders were asked to present their strategic plans and highlight areas of collaboration with UN Women while also taking the evaluation recommendations into consideration. The evaluation recommendations were thus incorporated in the design of UN Women continued work on gender equality and women's empowerment and reflected in the UNDAF 2014-2018.
-------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Case Study 11: UNDP - Assessment of Development Results in Uruguay

Full title	Assessment of Development Results of Uruguay UNDP Country Programme¹⁰
What were the objectives of the evaluation and the intended primary evaluation users?	The Assessment of Development Results (ADR), one main area of work of the Independent Evaluation Office (IEO) of UNDP, seeks to independently review progress of UNDP interventions nationwide. The 2014 ADR in Uruguay was led and conducted by the IEO, in collaboration with the Uruguayan International Cooperation Agency (AUCI). The intended primary users were the UNDP administration, UNDP Country Office and Regional Bureau for Latin America and the Caribbean, and national counterparts, including the government, civil society, donors, other UN agencies.
How was the evaluation used?	<p>In Uruguay, national counterparts expected that the evaluation would be an input for the next planning cycle of UNDP, but also as a contribution to a greater evaluation culture in a country in which public institutions are not commonly used to doing evaluations. Bilateral donors look at evaluation of the UNDP programme with interest. Limited evaluation culture in the government, however, presents important challenges and hinders the demand for evaluation by public institutions. The ADR exercise was considered very positive not only for the Country Office, but also for national counterparts. However, UNDP evaluation documents are perceived as too centred on UNDP, making it difficult to identify challenges common to the wider UN system.</p> <p>Evaluations have greater impact if there is follow-up, and should establish a roadmap to assess how and whether recommendations are taken on board.</p>

¹⁰ <http://web.undp.org/evaluation/evaluations/adr/uruguay.shtml>

	What is most valued, however, is the evaluation process itself. It is not so essential that the exercise be independent, but rather the rigor, and the quality of the process.
Why was the evaluation used and successful?	<p>There are a number of factors that contributed to the evaluation being used. The report clearly indicated factors that help projects achieve their goals; whether the programme adequately understood the context and issues to be addressed and whether strategic decisions were adequate; as well as the role of external factors. Also, it showed what has worked or not within sectors such as human rights or the environment. The report contained good and concrete recommendations and addressed changes in the external environment, such as a more restrictive financial environment or issues related to South-South cooperation. The two key elements are: 1) good evaluation design; and 2) understanding the context in which the programme being evaluated operates.</p> <p>It is also important to make a leap in the quality of project information for national counterparts, as in the case of Uruguay they are currently more used to activity reports than evaluations. There is need for indicators that are easy to understand and communicate and allow better monitoring. It is critical that quantitative information is available and that benchmarks are established. If the evaluation report is to be used by the government, as counterparts in constructing the next UNDP country programme, it must use language that is accessible and easy to understand for decision-makers. A good and short executive summary is extremely helpful. The report should also suggest how to make more efficient use of resources and align them with public policy as well as to how to engage with national counterparts and the civil society.</p>

Case Study 12: UNEP – Formative Evaluation of the UNEP Programme of Work

Full title	Formative evaluation of the UNEP Programme of Work (PoW) 2010-2011
What were the objectives of the evaluation and the intended primary evaluation users?	The UNEP Governing Council requested UNEP to prepare a Medium Term Strategy for 2010-2013 with clearly defined vision, objectives, priorities, impact measures and a robust mechanism for review by donors. Based on this strategy, UNEP developed biennial Programmes of Work (POWs) for 2010-2011 and 2012-2013, structured around six thematic sub-programmes consistent with the strategy. Expected accomplishments were defined for each sub-programme and PoW outputs specified for each of those expected accomplishments. This structure was the result of a reform process initiated by the current Executive Director and is leading UNEP towards the organization of its work around results-based priorities rather than divisional structures. The approach aims to improve coordination and reduce duplication of efforts, something that was perceived to be inherent to the previous situation.

	<p>The formative evaluation of the UNEP's PoW 2010-2011 intended to provide senior management with feedback on the design and delivery of the PoW. Specifically, the evaluation sought to understand whether projects were optimally linked to higher-level results. It also intended to provide feedback which may lead to adaptations in programme design and implementation and would, in turn, increase the likelihood of success in achieving the expected achievements and improve UNEP future planning processes.</p> <p>The evaluation was conducted as a desk study focusing on the processes and content of project/programme design and reporting arrangements in the PoW for 2010-2011. The evaluation was 'evidence-based', providing conclusions and recommendations based on objective and documented evidence to the extent possible. The evaluation approach involved the collection of qualitative and quantitative data from programme and project document reviews and interviews, and made extensive use of the 'Theory Based' approaches to examine project causality in UNEP's PoW.</p> <p>This evaluation was widely circulated and was discussed throughout the organization both at the preliminary findings stage and after the final conclusions and findings became available. Group discussions were organized with relevant stakeholders and comments were received from across the organization. Preliminary findings were shared with the Senior Management Team, and the final full report was presented in June 2011. The Executive Director of UNEP recommended the evaluation as 'essential reading' for all UNEP senior managers. Later that year, in November 2011, the Evaluation Office presented the findings during a retreat of the UNEP senior managers and regional directors, and at a meeting of the Committee of Permanent Representatives of the UNEP Governing Council. The report was disseminated by email, published on the website and printed. Methods used in the evaluation were the topic for a presentation at the UNEP EPE in Paris 2012.</p>
How was the evaluation used?	<p>Results from the formative evaluation were used by an in-house task team working on programme management and implementation. The team reinforced three major issues that were highlighted in the evaluation: 1) programme delivery in a results-based context; 2) resource allocation and alignment; and 3) accountability. The Evaluation Office tracked the progress of the organization towards the implementation of the recommendations and therefore the steps taken towards result-based planning and management, alignment of resources and better accountability. Out of the 21 recommendations, which were all pitched at a strategic level, 19 were closed as fully compliant and 2 as partially compliant.</p> <p>The formative evaluation has influenced strategic planning processes and is cited in major UNEP programming documents, for example the UNEP 2012-13 Programme of Work. Findings of the formative evaluation also featured in the</p>

	Executive Director's 2013 report to the Governing Council on the design of the 2014-17 medium term strategy The GA-approved UNEP Programme of Work 2014-15 also specifically cites key findings from the formative evaluation, as does the formal planning guidance issued by the UNEP Quality Assurance Section. The UNEP Programme Manual and in-house results-based management training incorporates many of the ideas promoted in the formative evaluation and is consistent with all of the recommendations for programme planning. This high-level usage was also confirmed in the UNEG Peer Review of UNEP.
Why was the evaluation used and successful?	The evaluation was strategically conducted at a time of profound organizational change. As it aimed to analyse such change and provide feedback, it received significant attention by relevant stakeholders, both at senior management level and organization-wide. The relevance of the findings ensured that its use was high and, as a result, it helped shape a new <i>modus operandi</i> within UNEP. For example, it led to a redefinition of the roles of sub-programme coordinators; it introduced a requirement to use a Theory of Change approach to project design; and it promoted better financial planning.

Case Study 13: UNEP – Midterm Evaluation of the Project for Ecosystem Services

Full title	Mid-term evaluation of the GEF-supported UNEP project “Project for Ecosystem Services”
What were the objectives of the evaluation and the intended primary evaluation users?	<p>The mid-term evaluation (MTE) of the Project for Ecosystem Services was conducted halfway through project implementation to analyse whether the project was on track, what problems or challenges the project was encountering and any corrective actions required. The project was implemented in five countries in three regions of the world and was considered a flagship project for UNEP on mainstreaming ecosystem services into development policy.</p> <p>The evaluation was conducted by an external consultant who was an expert in the project's subject field and had a sound background in evaluation but had no linkages to the evaluated project. The project was evaluated against standard evaluation criteria (relevance, performance, sustainability and factors affecting performance). It included a desk-based review of documentation, interviews, visit to UNEP headquarters to meet with the Evaluation Office and the project's implementation and execution teams, and visits to three of the five project countries. Prior to undertaking travels, the consultant prepared an Inception Report, including a Theory of Change of the project. The Theory of Change was then discussed and agreed upon with the project implementation and executing teams.</p>

	<p>The consultant adopted a proactive and participatory approach to the evaluation and engaged in in-depth discussions with the project implementation and execution teams and other stakeholders throughout the evaluation process. Evaluation findings were therefore communicated to the stakeholders throughout the evaluation. The consultant also participated in the project's second Steering Committee meeting, where she presented preliminary findings and recommendations of the evaluation. The completed evaluation report was sent by email to all project stakeholders and also published on the Evaluation Office website. The evaluation recommendations were compiled into a recommendation implementation plan, which was sent to the project implementing and executing teams for their action.</p>
How was the evaluation used?	<p>The evaluation provided a large number of recommendations ranging from general recommendations for the entire project to specific recommendations for each of the countries. The recommendations addressed a wide range of issues, related to communication, logframe, technical support, and technical quality checks. The recommendations were accepted by the project and an implementation plan was developed indicating what the project would do to address each of the recommendations, and by when. The team quickly revised the logframe according to the recommendations, helping the team to focus their attention on priority activities and outcomes. The Evaluation Office has been regularly reviewing the implementation plan and the recommendations are being adequately addressed.</p>
Why was the evaluation used and successful?	<p>The project team was satisfied with the consultant selection due to the consultant's experience and expertise. It is the view of the evaluation office that this increased the credibility of the evaluation and also helped to gain stakeholders buy-in to the evaluation, since the consultant engaged in in-depth discussions about the project and the subject field in general.</p> <p>The development and discussions of the project's Theory of Change prior to country visits was beneficial in terms of having a clear understanding of the project's logic prior to engaging in in-depth discussions with the stakeholders in project countries. Visits to some of the project countries were also viewed as highly beneficial for the evaluation since this enabled face-to-face discussions with stakeholders, helping to build a good relationship between the evaluator and the stakeholders and formulating useful recommendations. Participation in the Steering Committee meeting was highly beneficial in terms of making the SC members aware of the evaluation findings and gaining buy-in for the evaluation and its recommendations.</p>

Case Study 14: UNESCO - Evaluation of UNESCO's Standard-setting Work of the Culture Sector

Full title	Evaluation of UNESCO's Standard-setting Work of the Culture Sector Part I – 2003 Convention for the Safeguarding of the Intangible Cultural Heritage
What were the objectives of the evaluation and the intended primary evaluation users?	<p>The overall evaluation of UNESCO's Standard-setting Work of the Culture Sector comprised four evaluations, each focusing on one of UNESCO's main culture Conventions, and one audit of the working methods of the Conventions. Their purpose was to generate findings and recommendations regarding the relevance and effectiveness of the standard-setting work, with a focus on its impact on legislation, policies and strategies of Parties to the Conventions.</p> <p>The evaluation aimed to help UNESCO's Culture Sector, senior management and the Governing Bodies of the Conventions to strengthen, refocus and better coordinate the Organization's standard-setting activities. It also aimed to contribute to generating a better understanding of how the Conventions affect Parties' legislation and policies and the behaviour of key institutional actors.</p> <p>The primary users of this evaluation were the Governing Bodies and Parties to the Conventions, the UNESCO Convention Secretariat and Executive Board.</p>
How was the evaluation used?	<p>Following the finalization of the evaluation on the standard-setting work related to the 2003 Convention for the Safeguarding of the Intangible Cultural Heritage, the report was presented to the Intergovernmental Committee of the Convention (2013), which discussed and accepted the evaluation recommendations, and instructed the Secretariat and States Parties to take action accordingly. Further direction was subsequently given by the General Assembly of States Parties to the Convention (2014). The evaluation report, together with the other elements of the evaluation, was later presented to the UNESCO Executive Board (2014).</p> <p>The following are a few examples of how the evaluation was used (and is still being used). Several other activities are currently being implemented or planned for future implementation.</p> <ul style="list-style-type: none"> ○ Provision of concrete evidence of the workings and results of the standard-setting work related to the 2003 Convention; ○ Identification of good practices of successful implementation that are being used to improve knowledge sharing and learning by stakeholders; ○ Drafting of new operational directives on the contribution of intangible cultural heritage to sustainable development; ○ Establishment of one single mechanism for the assessment of nominations files to the various Convention mechanisms; ○ Substantive revision and improvement of UNESCO's capacity building programme in support of the implementation of the Convention

	<p>(including its content and format; the comprehensiveness of the training of trainers; the better diversification of the pool of trainers etc.). Steps have also been taken to improve follow-up on the capacity building activities;</p> <ul style="list-style-type: none"> ○ Revision of nomination and reporting forms to improve the evidence base of the results of the Convention, including its contribution to gender equality and women's empowerment; and its impact on policy and legislation. The revision will furthermore allow future NGO involvement in State Parties' Periodic Reporting on the implementation of the Convention; ○ Decision of UNESCO Executive Board to establish a working group to further discuss potential synergies and cross-cutting issues between the Conventions. ○ Use of the evaluation report as an important source of information and strategic guidance for decision making by the Convention Governing Bodies, the Convention Secretariat and by UNESCO staff working in field offices.
Why was the evaluation used and successful?	<p>The following factors contributed to ensuring that the evaluation is used and recommendations are implemented:</p> <ul style="list-style-type: none"> ○ Stakeholder involvement and extensive consultation during the entire process from evaluation design to finalization of the report; ○ Timing of the evaluation. After having been in force for seven years the time was ripe for a first assessment of the effectiveness of the Convention's mechanisms and of UNESCO's support activities, and for a stock-taking of the status and results of implementation. This created an opening for the implementation of the evaluation recommendations; ○ Demonstrated interest in the evaluation by the Governing Bodies of the Convention and (especially) the Convention Secretariat from the start of the exercise. This ensured that learning was happening throughout the evaluation process; ○ Increased donor expectations for UN Agencies to better demonstrate the results of their work and to improve its effectiveness and impact. ○ Novelty of this type of evaluation. Very few comprehensive evaluations of standard-setting work have been undertaken in the UN system so far. This evaluation therefore raised the interest of UNESCO Executive Board members and their readiness to support the implementation of the recommendations. ○ Quality of the evaluation recommendations, which were found to be relevant, sufficiently (but not overly) precise, and well-targeted. ○ Presentation of the report in various fora and meetings, and dissemination to a wide range of stakeholders not only through the evaluation office, but also through the Convention Secretariat and the expert community.

Case Study 15: UNESCO - Evaluation/ Review of UNESCO's Education Category I Institutes

Full title	Evaluation/ review of UNESCO's Education Category I Institutes
What were the objectives of the evaluation and the intended primary evaluation users?	<p>A review of six education-related institutes was carried out in collaboration with the UNESCO Education Sector. The aim of the review was to assess specific key aspects of the institutes' performance, their achievements and challenges within the context of the implementation of the overall strategy for those institutes and centres. In its findings and recommendations the review also considered the progress achieved in the implementation of the recommendations of the previous evaluation carried out in 2005/2006.</p> <p>The review was designed to inform the Education Sector's decision-making process towards strengthening the overall framework for cooperation with (and among) the institutes, as well as the strategic allocation of resources and human resource capacities. In terms of scope, the review covered the institutes' mandate and relevance, the results achieved, collaboration and interaction with partners within the UNESCO system and beyond, management and governance mechanisms, as well as different aspects of sustainability. The review resulted in individual reports for each institute, pointing to specific achievements and challenges of each, and directed to the institutes and their governing bodies. There was also a summary report that identified crosscutting issues and systemic recommendations, primarily directed to the Education Sector's senior management and UNESCO Member States, primarily the representatives of Members States in the UNESCO Executive Board.</p> <p>The primary intended users of this review were therefore the institutes' governing bodies and senior management, as well as the senior management of the UNESCO Education Sector and governing bodies. Furthermore, the individual recommendations for each institute provided a framework for internal reform and improvements concerning also programme and support staff of the institutes.</p>
How was the evaluation used?	<p>Systemic findings and recommendations pointed to the need for a clearer positioning of the institutes within the Education Sector's strategy, better demonstration and communication of results and more effective governance. These elements informed strategic reflection and decision-making within the Education Sector to improve strategic guidance, revisit the allocation of UNESCO resources to the institutes, and improve coordination and collaboration with the institutes. The Education Sector appreciated this review, which was carried out in parallel to an internal assessment on the management of the institutes. The review was instrumental in triggering a number reform proposals and decisions by governing bodies and senior management on necessary changes to the statutes of the institutes and</p>

	<p>establishing a framework for a more harmonized approach to collaboration, governance, management and operations.</p> <p>At the level of the individual institutes, the review provided a framework for reflection on how to better position each institute within the specific thematic area and institutional landscape in which it operated, as well as to initiate negotiations with the host countries to strengthen their support of the institutes. It also helped the institutes to implement internal reform efforts for better management, more focused and effective allocation of resources and for defining a stronger focus and results- based orientation of their work programmes.</p>
Why was the evaluation used and successful?	<p>Several factors contributed to the effective use of the review's findings (reflection and debate) and recommendations (implementation).</p> <ol style="list-style-type: none"> 1) At the sector level, the timing of the review was aligned to an internal assessment (complementary in terms of focus and scope), with both exercises feeding into the Education Sector's decision-making process on strategic and operational challenges for the future of the institutes. 2) The review generated credible and convincing evidence, which was presented in such a way as to coincide with the main strategic issues of the broader organizational reform process that was high on the agenda of UNESCO governing bodies. 3) The comparative aspect of the review was of particular interest to stakeholders as it informed reflections and decisions on resource allocation and systemic issues across institutes. 4) The review process was essentially a participatory process, in which extensive consultations with staff, management and governing bodies of individual institutes led to substantial ownership of the exercise by the different stakeholders. 5) Communication was effective – the individual reports were presented to the institutes' governing bodies and staff; the summary report was presented to the UNESCO Executive Board in the same session as the ED sector report on the institute management reform. The circulation of draft reports to different stakeholders was again an element that contributed to the overall sense of ownership and facilitated the subsequent process of implementation of the review's recommendations. 6) There was a systematic follow-up process over a period of two years. Periodic meetings between the stakeholders and evaluation officers on the basis of action plans, endorsed by all those involved, paved the way for implementing changes. 7) There was a general sense of appreciation by UNESCO governing bodies concerning the individual reviews and summary report as an input to decision-making.

Case Study 16: UNFPA – Joint Evaluation of the UNFPA-UNICEF Joint Programme on Female Genital Mutilation

Full title	UNFPA-UNICEF Joint Programme on Female Genital Mutilation: Accelerating Change
What were the objectives of the evaluation and the intended primary evaluation users?	The purpose of this evaluation was to assess the extent to which, and under what circumstances, the UNFPA-UNICEF Joint Programme accelerated the abandonment of female genital mutilation (FGM) in programme countries between 2008-2012. Besides serving as an accountability tool for programme countries, donors and other stakeholders, the evaluation was also envisaged as a learning opportunity. As such, it was intended to inform future UNFPA and UNICEF work on FGM.
How was the evaluation used?	<p>Following the finalization of the evaluation, a management response was jointly prepared by UNFPA and UNICEF senior management. Both the evaluation report and joint management response were presented to UNICEF and UNFPA Executive Boards.</p> <p>The evaluation recommended UNFPA and UNICEF to jointly work on a second phase of the programme. Based on the results and recommendations of the evaluation, the UNFPA/UNICEF coordination team worked on the design of the second phase. Out of the nine recommendations, three were specifically addressed with the second phase: (i) predictable longer-term financing, (ii) strengthening of the M&E system; and (iii) reinforcement of the regional level.</p>
Why was the evaluation used and successful?	<p>Several factors contributed to an influential and successful joint evaluation:</p> <ol style="list-style-type: none"> 1) This represented the first ever evaluation on the effects of FGM abandonment. 2) The evaluation was jointly managed by both Evaluation Offices (UNFPA Evaluation Office was the lead), with strong commitment and professional involvement of the two Evaluation Offices through the entire evaluation process. 3) The involvement of two senior, experienced evaluation professionals ensured quality and smooth delivery. 4) The evaluation had adequate funding and sufficient allocation of time for ToR preparation and the selection of the country case studies. 5) There was strong engagement by the joint Evaluation Reference Group (ERG). There was also clarity of roles for the joint Evaluation Management Group (EMG) and the joint ERG, and between the two.

	<p>6) The evaluation was based on a participatory approach including wide consultation with key stakeholders at global, regional, national and community levels, including final beneficiaries.</p> <p>7) There was optimal evaluation communication and use: the dissemination plan was developed jointly.</p> <p>8) There was a push from both agencies and strong collaboration to ensure a timely joint management response.</p>
--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Case Study 17: UNICEF - National Child Protection Agenda in Thailand

Full title	National Child Protection Agenda in Thailand
What were the objectives of the evaluation and the intended primary evaluation users?	<p>The evaluation was intended to provide inputs to the Royal Thai Government and UNICEF on how to strategically advance the national child protection agenda and strengthen the current national child protection system. The intended primary users of the evaluation were the Ministry of Social Development and Human Security (MSDHS) and other relevant ministries of the government, as well as the UNICEF Thailand Country Office (TCO). More specifically, the evaluation aimed to:</p> <ol style="list-style-type: none"> 1. assess the actual and potential contribution of child protection monitoring and response system (CPMRS) to the national child protection system; 2. determine the extent to which CPMRS had met its objectives; 3. determine the relevance, efficiency and sustainability of the CPMRS as an approach to strengthen the child protection system; and 4. provide recommendations for the refinement and potential scaling up of the CPMRS approach to the national level.
How was the evaluation used?	<p>Following the results of the evaluation, recommendations were presented to the MSDHS and UNICEF. This led to: (1) a change in child protection policies; and (2) UNICEF programme adjustments, both described below.</p> <p><u>Policy Change</u></p> <p>After several years of UNICEF advocacy with the Government for the expansion of the child protection monitoring and response system, the evaluation provided concrete evidence on the relevance and sustainability of the system. The evaluation process itself created interest among government partners and, at the end, led to wider discussions and agreements to support the implementation of CPMRS in other areas of Thailand. In 2013, as a result of the evaluation, the MSDHS expanded the system to selected districts in seven provinces, and they are planning to add more provinces this year. Moreover, the evaluation also led to the high level decision to develop the overall national child protection policy and strategy, which was also among its recommendations and which had been advocated by UNICEF for many years.</p> <p><u>Programme adjustments</u></p>

	The findings on effectiveness and efficiency of the CPMRS provided good guidance to the Government to embark on the expansion of the system. Based on the evaluation findings, UNICEF reviewed its guidelines on pilot projects and these were shared with other sections as a reminder of the important considerations and meaning of these guidelines.
Why was the evaluation used and successful?	<p>Several factors contributed to the impact of the evaluation:</p> <ol style="list-style-type: none"> 1) The evaluation met the interest of the audience, in this case government and non-government agencies dealing with children's rights. Indeed, there are still very few evaluations and studies available that take a holistic approach to child protection, and this evaluation therefore filled an important gap. It provided evidence that was timely and appropriate for the Thai Government to head toward the strengthening and improving child protection systems in the country. 2) The entire process of the evaluation involved the relevant stakeholders. The preparation of the terms of reference, design, methodology, timeframe and consultations involved policy makers at the national level to key beneficiaries at the community level. The analysis of results also underwent thorough consultations in order to make the evaluation comprehensive and ensure the buy-in and acceptance of key stakeholders. It is strongly believed that the quality involvement of key persons throughout facilitated the approval and adoption of the findings and recommendations. 3) The recommendations were found to be highly relevant, credible, feasible and well disseminated. They adequately addressed the underlying gaps that has long existed in the child protection work in Thailand and were perceived as highly credible. In order to improve the reach of recommendations, they were discussed with the Government in several meetings and translated into Thai.

Case Study 18: UNICEF - Review of the Global Education Cluster

Full title	Review of the Global Education Cluster Co-Leadership Arrangement between UNICEF and Save the Children (focus area 2)
What were the objectives of the evaluation and the intended primary	The creation of the Global Education Cluster co-leadership arrangement was a bold attempt to bring something new to the cluster approach to humanitarian response. It was – and remains – a unique organization of agency resources predicated on the expectation that a UN/NGO partnership might add value to the work of a cluster. As one of the last in a series of clusters to be created – and one that was controversial at the time – the Global Education Cluster offered an appropriate and timely vehicle for this pioneering experiment.

evaluation users?	This exercise constituted an independent evaluative review of the Global Education Cluster. It sought to identify and address gaps in the partnership, with the ultimate goal of improving sectoral coordination and achieving education results at field level. Recognizing that the exercise would not be relevant, credible, or used if undertaken by UNICEF alone, from the outset UNICEF sought to actively engage its evaluation counterparts at Save the Children to co-manage the review. This included joint finalization of the terms of reference, shared recruitment and day-to-day management of the consulting team, joint communications and co-chairing of reference group meetings, and regular co-manager meetings on strategic and technical issues.
How was the evaluation used?	A joint management response was subsequently issued, and led to the following key actions: a joint visioning exercise to articulate the cluster's objectives and indicators to monitor its performance; a roadmap for seizing on each partner's comparative advantage and a clarification of roles and responsibilities; a joint planning, budgeting as and resource mobilization process; stronger governance arrangements to help bridge inter-agency and single-agency accountabilities of the co-coordinators.
Why was the evaluation used and successful?	The close evaluation partnership was the single most pivotal factor that helped ensure that the review's recommendations were agreed and acted on by senior management of both organizations.

Case Study 19: UNICEF - Independent Review of UNICEF's Operational Response to the January 2010 Earthquake in Haiti (cross-cutting)

Full title	Independent review of UNICEF's operational response to the January 2010 earthquake in Haiti (cross-cutting)
What were the objectives of the evaluation and the intended primary evaluation users?	This exercise, commissioned by the Office of the Executive Director (OED), was titled a 'review' because of its focus on internal operational issues affecting UNICEF's corporate response to the Haiti earthquake, not on specific programmatic outcomes at field level. In this scenario, Haiti served to illustrate broader systemic challenges affecting UNICEF's ability to respond effectively to large-scale, sudden-onset and complex emergencies.

How was the evaluation used?	The evaluation and its management response, under OED leadership, has led to major changes to improve the organization's performance in large-scale emergencies. The most significant of these is the introduction of simplified standard operating procedures to more clearly guide UNICEF, as a highly decentralized organization, to respond to large-scale corporate emergencies. Other major actions include: greater integration of UNICEF's cluster work within its trainings and guidance for senior managers; fine-tuned human resources processes and systems for getting the right people on the ground at (and for) the appropriate time; clearer strengthened guidance to help UNICEF and its partners respond in urban disasters, and more. In addition, colleagues elsewhere in the organization reported applying relevant lessons from the Haiti review to emergencies in their own regions (Horn of Africa, Sahel).
Why was the evaluation used and successful?	<p>Two factors were pivotal in ensuring follow-through on the review. First, given the highly charged dynamics of the Haiti response, it was important that this independent review not only highlight critical shortcomings, but also demonstrate impartiality so as to garner credibility. This meant recognizing the many hurdles UNICEF faced in the Haiti response and avoiding 'finger-pointing' at individual corners of the organization, while still pinpointing systemic gaps. The impartial nature of the exercise formed a key talking point that was communicated to stakeholders throughout the exercise so as to pave the way for later acceptance of its critical findings.</p> <p>Second, as an OED-commissioned corporate exercise, it was critical that the Evaluation Office work hand in hand with OED throughout the exercise. Although key lessons were gained from this experience that will benefit future OED-commissioned exercises, OED's involvement efforts were vital to helping keep the review relevant to major policy currents in the organization, and sustain attention and positive engagement in the review.</p>

Case Study 20: UNICEF - Progress Evaluation of the Education in Emergencies and Post-Crisis Transition Programme

Full title	Progress evaluation of the Education in Emergencies and Post-Crisis Transition program
What were the objectives of the evaluation and the intended primary	The Education in Emergency and Post-crisis Transition (EEPCT) program aims to "put education in emergency and post-crisis transition countries on a viable path of sustainable progress toward quality basic education for all". The programme progress evaluation was finalized at the end of 2010. It examined global-level progress and entailed country case studies of Angola, Colombia, Cote d'Ivoire, Liberia, Philippines and Sri Lanka. A full management response followed and was presented to the Executive Board by special request in February 2012.

evaluation users?	<p>Funding for the programme was directed through well-established UNICEF channels, which track resources by donor, generating a good record of the flow of funds. While this mechanism provides for tracking of funds allocated and expended according to donor and country programme, it does not permit easy analysis of expenditure by goal or activity. The evaluation identified difficulties related to the flow of funds to countries. Funds are received late in the fiscal year, which leads to a scramble to allocate them in the year received. The process by which funds are allocated to countries was substantially improved in 2009 and 2010.</p> <p>Communication within UNICEF was not sufficient for country offices to understand EEPCT's aims and objectives. EEPCT has been used more as a fund to support existing country programmes than to support the programme's global objectives. In 2009, UNICEF undertook significant steps to address the lack of clarity regarding the objectives of EEPCT at the country level, and understanding at the country level has improved.</p>
How was the evaluation used?	<p>The management response workshop directly followed a presentation and Q&A by the evaluation team leaders, which provided workshop participants a comprehensive understanding of evaluation findings and recommendations. The workshop, led by the Education Section, included key stakeholders in headquarters. The evaluation team leader also participated in the workshop to answer questions and brainstorm possible actions. This provided an additional level of external insight, which was particularly appreciated by the education team.</p> <p>The evaluation served as a solid foundation for learning from past lessons for the development of the new Education and Peacebuilding Programme, funded by the Government of the Netherlands. The evaluation was cited in the new programme proposal, and evaluation findings on the weaknesses of past management structure, results frameworks, and harmonization of programme allocation and implementation were all considered in order to build a stronger future programme. The Evaluation Office played an integral role advising on the M&E plan and results framework in the new programme.</p>

Case Study 21: UNICEF - Joint Evaluation of the UNFPA-UNICEF Joint Programme on Female Genital Mutilation/Cutting (FGM/C): Accelerating Change

Full title	Joint evaluation of the UNFPA-UNICEF joint programme on female genital mutilation/cutting (FGM/C): Accelerating Change
What were the	The joint evaluation assessed the extent to which the UNFPA-UNICEF joint programme has accelerated the abandonment of female genital

objectives of the evaluation and the intended primary evaluation users?	<p>mutilation/cutting (FGM/C) in 15 programme countries (2008-2012). The evaluation provided an opportunity to ensure accountability to donors and other stakeholders, and was also a useful learning exercise. The evaluation assessed the relevance, efficiency, sustainability and the effectiveness of the holistic and multi-sectoral approach adopted by UNFPA and UNICEF in their programme for the acceleration of the abandonment of FGM/C. Furthermore, it assessed the quality of the coordination mechanisms established at the global level and within countries to maximize the effectiveness of joint programme interventions. Finally, the evaluation provided recommendations for the future direction of the FGM/C policies and programmes and gives UNFPA and UNICEF insights into the successes and challenges in conducting joint programming.</p> <p>This was the first collaborative evaluation between UNFPA and UNICEF evaluation offices that examined the relevance and effectiveness of the joint FGM/C programme. The evaluation was challenged to provide answers to questions related to the appropriateness of the approach used and results achieved. The evaluation had a clear utilization focus as the findings would be linked directly to decision-making on whether or not the Joint FGM/C Programme should go into second phase and what improvements needed to be made to improve programme effectiveness, expansion and sustainability but also efficiency of how UNFPA and UNICEF worked together as part of the joint initiative.</p>
How was the evaluation used?	The evaluation generated concrete evidence, lessons and recommendations, which were used for decisions related to the second phase of the joint programme. The evaluation validated the relevance and effectiveness of the social norms approach, which UNICEF is using with greater confidence for programming for addressing early marriage.

Case Study 22: UNICEF - Evaluation of UNICEF Emergency Preparedness Systems

Full title	2013 global evaluation of UNICEF emergency preparedness systems
What were the objectives of the evaluation and the intended primary	Emergencies have a negative effect on the realization of the rights of UNICEF core beneficiary groups. In 2012, UNICEF and its partners responded to 286 humanitarian situations of varying degrees in 79 countries. UNICEF involvement in emergency situations is expected to increase as emergencies become more frequent. It is, therefore, important that UNICEF effectively prepare for emergencies, both independently and in collaboration with national governments and partners, and also ensures that adequate investment has been made to this end. Recent audits and evaluations, however, have pointed to uneven emergency preparedness (EP) across emergencies.

evaluation users?	<p>This independent evaluation was commissioned by the UNICEF Evaluation Office (EO) to pinpoint the specific gaps in UNICEF EP policies and systems that need to be addressed or strengthened.</p> <p>This evaluation observed that UNICEF humanitarian activities are orientated towards emergency response rather than preparation or mitigation. Constraints identified in this report that hinder EP include: inadequate articulation of vision, goals, definitions and strategy; ad hoc funding; inconsistent application of programming; lack of integration of policies, practices and standards; limited accountability and lack of performance measurement and reporting activities.</p>
How was the evaluation used?	<p>Steps were identified to enhance UNICEF emergency preparedness in response to the Global Preparedness Evaluation, with a focus on more clearly linking preparedness and resilience. Immediate next steps include making emergency preparedness a part of annual work planning, reflecting preparedness in the PPP, and new modalities to fund preparedness.</p>

Case Study 23: UNICEF - Transforming Residential Institutions for Children and Developing Sustainable Alternatives

Full title	2011 Serbia: Transforming Residential Institutions for Children and Developing Sustainable Alternatives
What were the objectives of the evaluation and the intended primary evaluation users?	<p>In 2005, the Government of Serbia adopted the Social Welfare Development Strategy (SWDS), the main strategic framework for the reform of the social welfare system. One of the main goals of SWDS is “improvement of offer and quality of services in all forms of residential placement of beneficiaries”. The strategy foresees a decrease in the number of child placements in residential institutions and the introduction and application of new methodological approaches, new organization of work, and guaranteed quality of services adjusted, as far as possible, to beneficiary needs. The strategy envisages development of new services and service departments to support the life of children with disabilities or without parental care in the community, such as foster care, respite care, etc.</p> <p>The purpose of this evaluation was to: (1) evaluate the contribution of the project to the implementation of the Social Welfare Development Strategy, including contribution to the development of new policies and legislation in the area of child care; (2) identify approaches that were vital for the achievement of results as well as lessons learned and good practice examples that can become a knowledge base for future programming, and; (3) provide insight into the current status of the child care system and strategic</p>

	<p>recommendations for the next steps in the reform process relevant for all engaged stakeholders.</p> <p>Recommendations for MoLSP included: a set of actions aimed at improving management capacities and further planning of transformation of residential institutions within the MoLSP; provision of further support and education to residential institutions; establishment and capacity building of new Regional Centres for Fostering; strengthening of partnerships with local self-governments; and improved data about children.</p> <p>Recommendations for the Ministry of Health were related to: the creation of conditions for further and full implementation of Professional Methodological Guidance for Implementation of National Health Programme for Children, Youth and Women throughout Serbia in all maternity hospitals; further education of medical staff; and further strengthening of communication with parents of newborns at risk, with the aim to minimize institutionalization of such newborns.</p> <p>Special attention has been given to recommendations that might help finalize the Action Plan for the implementation of the Baby Friendly Health Initiative in hospitals.</p>
How was the evaluation used?	The evaluation recommendations were translated into action through a process of supporting the closure of large residential institutions for children and the transformation of 3 regular child-homes into small group-homes for children with disabilities, with possible transfer to foster families.

Case Study 24: UNICEF - Civil Registration Support in Cameroon

Full title	Civil Registration Support in Cameroon
What were the objectives of the evaluation and the intended primary evaluation users?	The country programme evaluated a decade of UNICEF interventions to increase access to birth registration and potential mechanisms for achieving sustainable results. The main conclusions of this evaluation report could be phrased in terms of more or less successful government, and UNICEF efforts to improve civil registration service delivery. The report's recommendation is to re-organize the civil registration service as a de-concentrated government service.
How was the	The key findings supported the programmatic decision to end direct services via the payment for birth certificates by UNICEF. Joint, sustained advocacy

evaluation used?	<p>with the donor community and faith-based organizations, alongside this evaluation of 10 years of UNICEF service delivery, led to the extension of the free birth registration period from 30 - 90 days, improving access as a part of civil status reform.</p> <p>Most notably, the evaluation of birth registration was linked to national civil status reform with recommendations for strengthening access included in the inter-ministerial road map. Similarly, the results of an assessment of alternative measures to detention were included in the review of the penal code and the new decree adopting such measures.</p> <p>Based on key findings, the UNICEF management response is to end financial contributions to the annual organization of the Children's Parliament, as the government provides funding for this purpose.</p>
-------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Case Study 25: UNRWA – Background Paper

Full title	Creation of a background paper on evaluation
What were the objectives of the evaluation and the intended primary evaluation users?	<p>The United Nations Relief and Works Agency for Palestine Refugees in the Near East (UNRWA) has delivered quite standardized services over a long period of time (for 65 years). As change is quite gradual there has been less need to develop documentation on change than in many other, faster paced environments. Documentation that ensures evaluability is therefore not as easily available as it would be in project environments that require logframes, theories of change and implementation plans. The lack of a clear theory of change and a common understanding of programme delivery throughout UNRWA can lead to misunderstandings during the evaluation process.</p> <p>To overcome this challenge and to improve evaluability, UNRWA introduced a process to create a background paper for each evaluation. The Evaluation Division spends time with the client to develop a background paper that includes the history of the programme, a theory of change, the scope, evaluation questions, and the objectives of the evaluation. The background paper is based on analysis of relevant documentation and discussions with key stakeholders. During this process the Evaluation Division typically moderates a meeting of stakeholders to come to an agreement on the theory of change of the programme before the background paper is finalized. Drafts of the background paper are shared with primary stakeholders to correct factual errors and to confirm the theory of change. The background paper is later endorsed by the evaluation steering committee.</p>

How was the evaluation used?	Since the introduction of the background paper, evaluability has greatly improved as the theory of change forms the basis to evaluate the effectiveness and impact of the intervention. Collaboration with the evaluation client has also greatly improved as the scope and evaluation questions are discussed and endorsed, so there is less chance for misunderstanding as the evaluation progresses.
Why was the evaluation used and successful?	The Evaluation Division experimented with developing the background paper using consultants or in-house. The development of the background paper in-house has been much more efficient, as staff better understand the context of UNRWA and are able to have more frank discussions with the clients.

Case Study 26: UNRWA – Evaluation Steering Committees

Full title	Creation of a steering committee for each evaluation
What were the objectives of the evaluation and the intended primary evaluation users?	<p>The United Nations Relief and Works Agency for Palestine Refugees in the Near East (UNRWA) has many stakeholders that are interested to use evaluations, but would typically not get closely involved in the process. At the same time, although evaluations are public, few people use them to reflect about the programs.</p> <p>To get key stakeholders involved and publish evaluations for use, UNRWA as a standard creates steering committees with internal and external stakeholders to guide the evaluation process.</p>
How was the evaluation used?	The steering committee approves the background paper for the evaluation, discusses the inception report, is part of the presentation of the preliminary findings and recommendations and provides comments on the draft report.
Why was the evaluation used and successful?	The involvement of the steering committee has led to the evaluations being used by government and donors as well as by UNRWA internal stakeholders. Discussing findings in a forum of senior stakeholders greatly increases the probability of the findings being used and the recommendations being implemented.

Case Study 27: UNRWA – Interactive recommendation follow up

Full title	Interactive recommendation follow up
What were the objectives of the	The Evaluation Division of the United Nations Relief and Works Agency for Palestine Refugees in the Near East (UNRWA) had previously followed up recommendations using a model from the Audit Division - communicating to the owners of the recommendations that an update towards the

evaluation and the intended primary evaluation users?	<p>implementation of recommendations is due at a certain date. Unfortunately, this formal follow up has not resulted in a high response rate.</p> <p>To address this challenge, the Evaluation Division added some interactive components to the recommendation follow up, including dialogue with clients and meetings to discuss the recommendation follow up progress.</p>
How was the evaluation used?	During the meetings with clients the progress on the different recommendations is discussed. This allows the Evaluation Division to reach out to clients and clarify what evaluation is in comparison to other oversight instruments and advocate for the implementation of the recommendations. The discussions clarify the spirit of the recommendations and what actually would be required to close the recommendations.
Why was the evaluation used and successful?	Clients have an opportunity to reflect on their programs while engaging on the nature of the recommendations. This enables them to step back and spend some time on strategic thinking. As a result of this process there is now a 100 per cent response rate on the recommendation follow up process and a greatly improved implementation rate for recommendations.

Case Study 28: UNRWA – Evaluation of Agency Medium Term Strategy

Full title	Evaluation of the UNRWA medium-term strategy
What were the objectives of the evaluation and the intended primary evaluation users?	<p>In 2009, the United Nations Relief and Works Agency for Palestine Refugees in the Near East (UNRWA) created its first medium-term strategy. This strategy guides UNRWA service delivery by defining four in the five fields of UNRWA service delivery (A long and healthy life, acquired knowledge and skills, a decent standard of living, and human rights enjoyed to the fullest extent possible).</p> <p>During discussions with host and donor governments it became clear that a mid-term evaluation of the medium term strategy would be very much appreciated to guide the development of the subsequent strategy.</p>
How was the evaluation used?	The evaluation used a steering committee with participation of a host government and donor governments ensuring high visibility of the process. The mid-term evaluation of the medium term strategy was discussed in the Advisory Committee of UNRWA and UNRWA committed to implement the recommendations from this evaluation.
Why was the evaluation used and successful?	The interest of host and donor governments and the timeliness of the evaluation enabled the use of the evaluation. As a result the next medium term strategy was developed in a more participatory manner. In addition the monitoring and evaluation framework is integrated in the new medium term

	strategy with evaluations aligned to the strategy's strategic areas over the coming six years.
--	------------------------------------------------------------------------------------------------

Case Study 29: WFP - Transition from Food Aid to Food Assistance

Full title	A synthesis of four strategic evaluations on the transition from food aid to food assistance
What were the objectives of the evaluation and the intended primary evaluation users?	A 2012 synthesis carried out by the WFP Office of Evaluation (OEV) drew from four independent strategic evaluations that assessed different aspects of WFP transition from a food aid to a food assistance agency, as called for in the Strategic Plan 2008-2013. The individual evaluations covered WFP role in social protection and safety nets; ending long-term hunger; working in partnership; and how country offices adapt to change. The synthesis revealed striking similarities in the findings, conclusions and recommendations of the four evaluations, concluding that organizational support for the transition was weak and that the adaptation of necessary systems, procedures and staff capacities lagged behind the pace of change in the field.
How was the evaluation used?	The synthesis informed WFP policy and operational development, particularly in regards to enhancing WFP capacity to advise and support governments, to raise funds and engage in partnerships. The synthesis was heavily referenced in the Strategic Plan 2014-2017. For example in response to evaluation findings, the new strategic plan placed a greater emphasis on how strategic shifts were to be achieved. The synthesis was the most frequently cited in the recent Office of Evaluation Peer Review as influential by Board members, senior management, regional and country based staff, and OEV itself. The Office of Evaluation noted that senior management's attention to the synthesis was unprecedented. The synthesis, the individual evaluation "Working in Partnership" and a related strategic evaluation of WFP's Private Sector Strategy also informed the development of WFP's first Partnership Strategy.
Why was the evaluation used and successful?	The use of the evaluation results was most likely due to the importance that WFP placed on transitioning the organization to a food assistance role and the attention paid to the transition by the WFP Executive Director and senior management with strong interest of the WFP Governing Body. This occasioned widespread support – both internal and external – for the sorts of changes needed to more effectively bring about the shift. The synthesis was particularly helpful in pointing out the tools and operating principles that are needed at all levels to bring about the change. The report's conclusions on the practical requirements for organizational change – capacity, funding, technical support, and partnerships – enabled the new strategic plan to take a more realistic approach to achieving its objectives. Importantly, many of the constraints identified by the synthesis were internal and systemic and thus within WFP

	control to address. This made it possible for the organization to address practical suggestions concerning leadership, guidance, human resources development and partnering strategies.
--	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Case Study 30: WFP- Food Assistance in Bangladesh

Full title	The contribution of food assistance to durable solutions in protracted refugee situations; its impact and role in Bangladesh: a mixed method impact evaluation
What were the objectives of the evaluation and the intended primary evaluation users?	Since 1991, WFP has assisted approximately 30,000 Rohingya refugees from Myanmar living in two refugee camps in Cox's Bazar district, Bangladesh. The refugees are dependent on humanitarian assistance as official regulations restrict their movement outside the camps and involvement in income-generating activities. In 2010-2011, an evaluation was jointly commissioned by WFP and the Office of the United Nations High Commissioner for Refugees (UNHCR) to assess the role of food assistance in, and its contribution to, self-reliance and durable solutions for the refugee and the refugee-affected populations.
How was the evaluation used?	<p>The evaluation found that food assistance delivers short-term food security outcomes, but that overall dietary diversity is poor among refugees. Furthermore, assistance should adapt to the protracted context, within an overall transition strategy. To this end, the evaluation recommended that alternative food assistance mechanisms be developed that enable more accurate targeting, are more appropriate to the refugee's livelihood reality, and are more cost effective and efficient.</p> <p>WFP's Bangladesh Country Office team sought a solution that provides the refugees with greater choice of food items to cover daily nutrition needs. Vouchers were identified as a more viable option than cash due to concerns that cash grants might encourage increased migration of Rohingyas from Myanmar. The Government of Bangladesh approved food distribution through e-vouchers in December 2013 and WFP launched the vouchers in 2014.</p> <p>The vouchers enable refugees to purchase a range of food items according to their families' needs and preference from contracted vendors inside the refugee camps. Women cardholders, whose customs constrain them from moving outside the camps, are able to safely access food from these shops.</p> <p>The e-voucher modality is less costly to deliver than food distribution and gives refugees greater choice in the foods they consume and when they can access their entitlements. E-vouchers are expected to enhance the nutritional value of the assistance, increase the security and accuracy of the assistance and</p>

	have positive 'spill-over effects' on the domestic economy by providing business to food traders and shops. Registration is streamlined and the e-voucher has the potential for UNHCR to include non-food items in the future.
Why was the evaluation used and successful?	The evaluation recommendations were reinforced by several other internal and external assessments, and were supported by the strong demand within the refugee community for a new modality. Bangladesh officials visited Turkey where government officials demonstrated the use of cards with Syrian refugees, which helped increase the confidence of Bangladesh officials in the use of e-vouchers. Strong leadership was demonstrated by the WFP Country Office senior management, and support provided by various technical units in WFP. The partner UNHCR, with support of the Bangladesh Government, provides a database of refugees and support for beneficiary registration.

Case Study 31: WFP – School Feeding in the Gambia

Full title	School Feeding in The Gambia (2001-2010): A mixed method impact evaluation
What were the objectives of the evaluation and the intended primary evaluation users?	Between 2001 and 2010, WFP supported three school feeding programmes in The Gambia. The programmes reached 113,000 rural students each year, about 40% of all primary school children in the country. A 2011 impact evaluation of the WFP school feeding programme in The Gambia took place at a time when the Gambian government was considering new directions for the school feeding programme, based on the assumptions of long term impact.
How was the evaluation used?	<p>The Gambian government welcomed the evaluation, which, for the first time, provided rigorous impact information about WFP school feeding. The evaluation found clear evidence that school feeding contributed to the nutritional requirements of participating students. However, the quality of education and certain practices that excluded some children from the programme tended to limit the impact of school feeding on net enrolment.</p> <p>In response to the evaluation recommendations, the Gambian government established an inter-sectoral task force to coordinate and monitor the National School Feeding Programme and Policy. WFP enhanced its technical assistance and capacity support to enable an eventual handover of school feeding to the government. WFP and the government developed an improved targeting system for school feeding. The impact evaluation has been cited in government policies, such as the Millennium Development Goals Accelerated Framework, the revised</p>

	Education Policy (2011-2012) and The Gambia's Programme for Accelerated Growth and Employment. It was also integrated into a national level workshop on the future of school feeding in The Gambia.
Why was the evaluation used and successful ?	Access to education has been a priority for The Gambia since the achievement of independence in 1965. Girl's education, in particular, has received significant attention and has been the object of national and international efforts since the 1990s. Starting in 2000, The Gambia's nutrition policy has prioritized improving the nutrition of women and school-age children. The Ministry of Basic and Secondary Education has identified the quality of education as a major priority and is working to improve monitoring and teacher performance. The government's interest in extending and taking control of the school feeding programme grew from these commitments and was predicated on a better understanding of the long-term benefits, which was part of WFP rationale to conduct the impact evaluation in The Gambia. The evaluation findings and practical recommendations reinforced the government's commitment and enabled The Gambia to move ahead with its plans for a national school feeding programme, with WFP support.

Case Study 32: WIPO – Knowledge Sharing Evaluation

Full title	Knowledge-sharing evaluation in WIPO
What were the objectives of the evaluation and the intended primary evaluation users?	The 2014 IOD evaluation of knowledge sharing in WIPO was strategically conducted to reinforce the Joint Inspection Unit (JIU) Management and Administration Review report on the need for a structured approach for knowledge management within WIPO.
How was the evaluation used?	This joint initiative contributed to the identification of the main assets and needs in knowledge sharing, enabling the evaluation to provide a sequence of recommendations under the form of a roadmap that addressed the various stages to implement a knowledge sharing strategy.
Why was the evaluation used and successful?	The inclusion in the evaluation of a Learning Resource Group, and the collaborative work with WIPO administration and staff, were essential in gathering all perspectives, and raising widespread awareness of the strategic opportunities and benefits that adequate knowledge sharing procedures bring to the Organization.

	<p>For the first time, the administration has requested the evaluation to continue its support by assisting in the process of defining a management action plan based on the evaluation recommendations. By so doing, the evaluation is still contributing to ensure the adequate understanding and appropriate allocation of resources to finalize the comprehensive knowledge sharing strategy, as per indicated by the JIU.</p>
--	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Case Study 33: WIPO – Recommendations from IOD Evaluation Reports

Full title	Recommendations from Internal Oversight Division evaluation reports
What were the objectives of the evaluation and the intended primary evaluation users?	<p>Recommendations from Internal Oversight Division (IOD) evaluation reports have been utilized to strategically plan further WIPO interventions with Member States. A significant example is the implemented recommendations of a 2013 evaluation on the support services internally provided to the WIPO Intergovernmental Committee on IP and Genetic Resources, Traditional Knowledge and Folklore. As reported, the early implementation of these recommendations has contributed to provide better information to participants, to enhance communicating processes, and to provide better support to Member States during the preparation of upcoming sessions.</p>
How was the evaluation used?	<p>Development Agenda project evaluations were also covered by IOD's early stages of work and have had a reported impact on WIPO efforts in this area. A significant case was the use of results towards the establishment of a sustainability-led transition plan jointly developed with the beneficiary Member State, and the use by the program of specific methodological tools and frameworks.</p>
Why was the evaluation used and successful?	<p>Through the IOD 'one-year after' questionnaire, users identified key factors for actionable recommendations. with extremely professional conduct of the evaluation and the extent to which it addressed issues relevant for the unit.</p> <p>Ex-post satisfaction surveys to the evaluation clients rated recommendations as highly actionable. Including relevant areas for the programme, along with objective and timely delivery of results, in the scope of the evaluation was essential for the positive impact on the management processes and the achievement of the objectives of the programme.</p>