Inspection and Evaluation Manual

United Nations Office of Internal Oversight Services
Inspection and Evaluation Division (OIOS-IED)

OIOS
Office of Internal Oversight Services

Start

How to use this manual

This is an interactive manual which allows the user to quickly link to other parts of the document and to other information sources. For example, from the table of contents you can move directly to a section of interest to you by simply clicking on the section header. The following links are provided:

↑ Table of Contents	Takes you back to the Table of Contents
↗ Top of this section	Takes you to the beginning of the section you are in
Part I, Section 1.4.	Takes you to another part of the manual
[IED #2]	Takes you to OIOS-IED's intranet, which contains links to a number of internal OIOS-IED resources
[➔ United Nations. 2000a]	Takes you to the OIOS-IED website, which contains links to a number of external resources
[📖 UN Chart]	Takes you directly to to external websites

This manual is best viewed with Adobe Reader.
You can download the latest version free at
http://get.adobe.com/uk/reader/

© United Nations Office of Internal Oversight Services (OIOS), Inspection and Evaluation Division (IED), New York, September 2014

Image cover page (Blue Globe Puzzle): 123RF Limited

This manual contains links to many resources, some of which are publically available and others of which are internal to the OIOS-IED and therefore proprietary. OIOS-IED's internal documents are located on its server and are regularly updated. The links to these internal documents are therefore not immediately accessible to external users of this manual. Readers wishing to obtain the latest version of an internal OIOS-IED document are kindly requested to email Ms. Catherine Nyawire (nyawire@un.org), and OIOS-IED will do its utmost to accommodate such requests.

Foreword

On behalf of the Office of Internal Oversight Services Inspection and Evaluation Division (OIOS-IED), I am pleased to share with you this Inspection and Evaluation Manual. The result of a collaborative process drawing on the Division's own experience, as well as a wealth of external sources, it represents a revision to OIOS-IED's first manual, developed shortly after the Division's formal establishment in January 2008.

OIOS-IED's mandate remains unchanged since 2008. Today as before, the Division produces independent inspections and evaluations of UN Secretariat programmes' effectiveness (and, wherever feasible, impact), relevance and efficiency on behalf of the Secretary-General and Member States. Toward this end, OIOS-IED remains firmly committed to providing timely, relevant, objective and credible information that its stakeholders – including programme managers themselves – can use to strengthen the Organization's performance. OIOS-IED is guided in its work by the United Nations Evaluation Group (UNEG) norms and standards.

Also unchanged is OIOS-IED's role in relation to other entities within the UN system. Within OIOS itself, its role is differentiated from that of the Office's two other oversight functions: audits, which focus on internal controls and compliance with UN rules and regulations, and investigations, which focus on the determination of misconduct and wrongdoing. And in contrast to the self-evaluation functions embedded within Secretariat programmes themselves, OIOS-IED is operationally independent from the programmes it evaluates.

Finally, the objectives and purpose of this revised manual remain identical to those of its predecessor, namely to provide clear guidance to OIOS-IED staff, ensuring consistency in quality standards as well as processes and procedures, that help the Division achieve its ultimate goal: producing credible inspections and evaluations that make a difference.

Despite such considerable continuity, much has changed in the six years since OIOS-IED released its first manual. Since its formal establishment, OIOS-IED has sharpened its strategic focus. This process led in 2012-13 to the development of OIOS-IED's Programme Impact Pathway (PIP), a logic model or visual "roadmap" conveying what the Division aims to achieve and how it aims to achieve it. Likewise, OIOS-IED's experience navigating the complexities of programme evaluation in the UN Secretariat has expanded, enabling its evaluators to glean lessons and good practice from each other, drawing on a much larger body of first-hand knowledge than was previously the case.
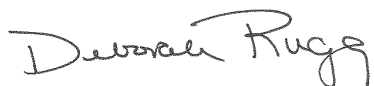
At the same time, in a climate marked by significant resource constraints, OIOS-IED evaluators have had to adapt and become ever nimbler, ever more flexible and ever more creative in their approaches to solving the specific challenges that each new inspection or evaluation presents. This need has prompted internal demand for more, and more concrete and custom-tailored, guidance to help staff produce the most timely, relevant, objective and credible reports that can make a difference – and to do so in the most efficient way possible. Fortunately, the toolkit of technologies and methods available to evaluators has grown during this same period, affording OIOS-IED staff access to a wider range of fit-for-purpose external supports than ever before.

Lastly, along with OIOS-IED's evolution has come a more explicit focus on its relation to other actors. For example, while OIOS-IED has always viewed its relationship with programme managers as a key aspect of its work, the present manual offers more specific guidance on how to navigate this relationship – and how to do so without compromising its independence – to help maximize evaluations' success. Similarly, while OIOS-IED has always been a contributor to strengthening the self-evaluation function within the UN Secretariat, this revised manual was developed with an explicit eye to helping achieve this objective, even though the manual's primary audience remains OIOS-IED's own staff.

With these changes in mind, this revised manual offers five main areas of improvement over the original. These include:

1. A firm grounding in OIOS-IED's targeted results, as summarized in its PIP;
2. Additional points of guidance, including both new areas in which guidance is provided, as well as expanded and more concrete points of guidance in areas already covered in the previous version of the manual;
3. Expanded resources, including updated examples of OIOS-IED 's own good practice and a significantly expanded list of external resources;
4. A more interactive and user-friendly format, including internal links within the document as well as hyperlinks to the internal and external resources described above, as well as greater use of graphics to visually illustrate key concepts; and
5. A wider target audience, with wider applicability to colleagues in the UN Secretariat self-evaluation function and programme management, in addition to the manual's main audience, OIOS-IED evaluators.

I hope this revised manual proves useful in your work to bring evaluation knowledge to bear on programme performance. OIOS-IED will continue to periodically review and update the manual to accommodate new developments. We welcome your feedback as you use it!

Deborah Rugg, Director
Office of Internal Oversight Services
Inspection and Evaluation Division (OIOS-IED)
Chair, United Nations Evaluation Group (UNEG)

September 2014

Table of contents

Part I.   About OIOS-IED

Part I of this manual discusses the foundations of the Inspection and Evaluation Division of the Office of Internal Oversight Services (OIOS-IED) - what it is, what it does, and the normative principles that shape its work - within the United Nations (UN) Secretariat and the wider UN system of which it is a part.

Toward this end, this Part provides an overview of the mandates underlying OIOS-IED's work, the core principles that define the Division's approach to its work, the operational and human resource structures within which it operates, and the work planning and quality assurance processes it relies on to ensure that this work leads to the Division's ultimate aim: providing timely, objective, credible and relevant inspections and evaluations (Part I, Section 1.1) that are used to improve UN Secretariat programmes' performance, and by extension the achievement of the objectives of the UN.

Together, these foundations underpin OIOS-IED's vision of being the best source of information on whether or not the UN works well. OIOS-IED does this through its mission to produce "world-class inspections and evaluations, based on the highest standards of oversight professionalism, that assist the UN in becoming the most efficient and effective Organization possible and to support it in reaching the objectives, ideals and aspirations embodied in the Charter."

This part of the OIOS-IED manual, Part I, has three chapters:

• OIOS-IED's Work - what OIOS-IED does and with what ultimate purpose in mind, and the resources and normative principles it relies on to achieve its targeted results (Part I, Chapter 1);

• OIOS-IED's Approach to Work Planning - how OIOS-IED plans its inspections and evaluations in order to fulfil its key stakeholders' expectations, and to do so in line with its core principles (Part I, Chapter 2); and

• The OIOS-IED Quality Assurance System (QAS) - how OIOS-IED goes about ensuring that its inspections and evaluations are of high quality so that they are credible, and therefore actionable (Part I, Chapter 3).

1.   OIOS-IED's Work

This chapter focuses on what OIOS-IED does and with what ultimate purpose in mind, and the resources and normative principles it relies on to achieve its targeted results. It has seven main sections:

• OIOS and IED Mandates (Part I, Section 1.1)
• The OIOS-IED Programme Impact Pathway (PIP) (Part I, Section 1.2)
• OIOS-IED's Oversight Universe (Part I, Section 1.3)
• OIOS-IED Products (Part I, Section 1.4)
• UNEG Norms and Standards (Part I, Section 1.5)
• Ensuring Human Rights and Gender-sensitive Evaluations (Part I, Section 1.6)
• OIOS-IED Staff and Financial Resources (Part I, Section 1.7)

1.1   OIOS and IED Mandates

The Office of Internal Oversight Services (OIOS) was established in 1994, under General Assembly (GA) resolution 48/218 B of 29 July 1994, to enhance internal oversight responsibilities in respect of the resources and staff of the United Nations (UN). Its internal oversight mandate extends to the entire UN Secretariat, but does not include the UN funds, programmes or specialized agencies.

The main location of OIOS is at the UN Headquarters in New York. Its Investigations Division (ID) and Internal Audit Division (IAD) have representations in other UN Headquarters locations, including Nairobi, Geneva and Vienna, as well as in various peacekeeping operations (PKOs).

OIOS is led by the OIOS Under-Secretary-General (USG). The USG is appointed by the UN Secretary-General (S-G) following consultations with UN Member States, and approved by the GA for one five-year term without possibility of renewal.

OIOS's operational independence is critical in order for it to be able to carry out its mandate effectively. GA resolution 48/218 B provides the original legislative basis for its operational independence: "The Office of Internal Oversight Services shall exercise operational independence under the authority of the S-G in the conduct of its duties and, in accordance with Article 97 of the Charter, have the authority to initiate, carry out and report on any action which it considers necessary to fulfil its responsibilities with regard to monitoring, internal audit, inspection and evaluation and investigations as set forth in the present resolution." OIOS's exercise of its independence is further elaborated in ST/SGB/273 of 7 September 1994.

OIOS MANDATE

In the original resolution establishing OIOS, the GA mandated OIOS to conduct internal audits, inspections and evaluations, as well as investigations into reports of violations of UN rules, regulations and pertinent administrative issuances. It also called for the GA Fifth Committee to regularly review the functions and reporting procedures of OIOS. This original mandate has led to a number of further provisions on OIOS over the years (Table 1).

To carry out its oversight mandate, OIOS is organized into three divisions. The Internal Audit Division (IAD) is responsible for the audit function, the Investigations Division (ID) for the investigation function, and the Inspection and Evaluation Division (IED) is responsible for the inspection and evaluation functions.

How the Office and its three divisions implement the OIOS mandate, and how they are required to report to the GA on their performance, are defined in the OIOS Strategic Framework [IED #1]. Approved by the GA, the Strategic Framework is the official document UN Member States use to set expectations on the overall results they want OIOS to achieve - i.e., timely, objective, credible and relevant oversight results that make a difference - along with the indicators they want OIOS to use in its performance reporting [📖 OIOS Annual Reports] against these targeted results. It also includes a brief summary of OIOS's overall strategies for achieving these results. The OIOS Strategic Framework covers a two-year period, or biennium.

Complementing the Strategic Framework is the OIOS Programme Impact Pathway (PIP) [IED #2], a visual tool that conveys OIOS's planned "roadmap to results" - what it aims to achieve as set forth in the Strategic Framework, and the inputs, activities and outputs through which it aims to achieve these results. In addition to the OIOS PIP, each of the Office's three divisions has developed a PIP that specifically depicts the results it seeks to achieve within the broader OIOS mandate and Strategic Framework, and how it aims to do so.

Figure 1 provides an overview of OIOS's three divisions and what each contributes to the Office's broad oversight objective.

Figure 1: OIOS Divisions and Their Complementary Oversight Functions

As Figure 1 suggests, although each of OIOS's oversight functions is distinct, all three are also mutually complementary. This complementarity is depicted in the figure by a small centre area of three-way overlap. Collectively, OIOS's three divisions seek to contribute to the relevance, effectiveness, impact and efficiency of UN Secretariat programmes and to increase their transparency and accountability.

OIOS's three divisions are also complementary to each other in practical ways. They often coordinate their activities - e.g., by sharing their work plans and engaging in ad hoc communication in order to share relevant information and avoid unnecessary overlap or duplication. Beginning in 2014, the Divisions are piloting the notion of a joint project, in which staff members collaborate across divisions, bringing their respective expertise to the examination of a programme or theme, so as to maximize complementarity and potentially achieve greater efficiency than single-division projects sometimes entail.

These areas of inter-divisional coordination are depicted in Figure 1 by the areas of two-way overlap. The importance of such coordination, and of coordination with relevant external stakeholders, is also reflected in OIOS's divisional PIPs.

OIOS-IED MANDATE

OIOS-IED was formally established on 1 January 2008 after deliberations by UN Member States and the UN Secretariat in the context of the 2005 World Summit-mandated "Comprehensive Review of Governance and Oversight within the UN and its Funds, Programmes and Specialized Agencies" (A/RES/61/245) (Table 1). Prior to 2008, OIOS-IED had been known as the Monitoring, Evaluation and Consulting Division (MECD).

OIOS-IED conducts independent inspections and evaluations of individual UN Secretariat departments and programmes on behalf of the S-G and Member States; it also undertakes inspections and evaluations of topics that cut across multiple UN Secretariat departments or programmes; these two types of exercises are referred to as programme evaluations and thematic evaluations, respectively. (Part I, Section 1.1).

The specific mandate for the OIOS-IED inspection function is articulated in ST/SGB/273 (Table 1). It states that OIOS "shall conduct ad hoc inspections of programme and organizational units whenever there are sufficient reasons to believe that programme oversight is ineffective and that the potential for the non-attainment of the objectives and the waste of resources is great, and otherwise as the USG for Internal Oversight Services deems appropriate. These

inspections shall recommend to management corrective measures and adjustments as appropriate."

The mandate for OIOS-IED's evaluation function originates in GA resolution 37/234, which was later re-affirmed and expanded in scope of detail by GA resolution 48/218 B and further elaborated in ST/SGB/2000/8 (Table 1). ST/SGB/2000/8 defined the objective of evaluation as being:

- To determine as systematically and objectively as possible the relevance, effectiveness, impact and efficiency (evaluation criteria) of the Organization's activities in relation to their objectives; and

- To enable the Secretariat and Member States to engage in systematic reflection, with a view to increasing the effectiveness of the main programmes of the Organization by altering their content and, if necessary, reviewing their objectives.

Table 1 provides an overview of the main sources of OIOS-IED's mandate.

In practice, OIOS-IED's inspections and evaluations are similar in nature, differing mainly in their scope and duration. Therefore, for ease of comprehension, throughout the remainder of this manual the term "evaluation" will encompass both inspections and evaluations. The term typically used for a UN Secretariat department or programme being evaluated is the evaluand.

Table 1: Main Sources of OIOS-IED's Mandate

A/37/38 [➔ United Nations. 1982a] of 9 June 1982 entitled Report of the Committee for Programme and Coordination on the work of its twenty-second session	Contains the mandate for the UN Committee for Programme and Coordination (CPC) to review the implementation of its recommendations three years after taking decisions on evaluations submitted to the Committee, thus mandating OIOS-IED to conduct Triennial Reviews (Part I, Section 1.4).
A/RES/37/234 [➔ United Nations. 1982b] of 21 December 1982 entitled Programme planning	Contains in Article 6 the original objective of and mandate for evaluation.
A/RES/48/218 B [➔ United Nations. 1994a] of 29 July 1994 entitled Review of the efficiency of the administrative and financial functioning of the United Nations	Reaffirms and expands A/RES/37/234. Establishes OIOS to enhance the oversight functions within the UN Secretariat and provides the legislation for its operational independence.
ST/AI/397 [➔ United Nations. 1994b] of 7 September 1994 regarding Reporting of inappropriate use of United Nations resources and proposals for improvement of Programme delivery	Informs staff and others of the procedures for confidential reporting of possible misuse of funds, waste or abuse of UN facilities or privileges, and/or for making proposals for the improvement of Programme delivery.
ST/SGB/273 [➔ United Nations. 1994c] of 7 September 1994 entitled Establishment of the Office of Internal Oversight Services	Elaborates on the operational independence of OIOS, and further articulates the specific mandate for the OIOS-IED inspection function.
ST/AI/401 [➔ United Nations. 1995a] of 18 January 1995 entitled Personnel arrangement for the Office of Internal Oversight Services	Outlines the administrative arrangements and authority of the OIOS USG in personnel matters, and further underlines OIOS's operational independence.
ST/SGB/1997/5 [➔ United Nations. 1997a] of 12 September 1997 entitled Organization of the Secretariat of the United Nations	Establishes the organizational structure of the UN Secretariat.

ST/SGB/1998/2 [➔ United Nations. 1998a] of 12 February 1998 entitled Organization of the Office of Internal Oversight Services	Establishes OIOS's organizational structure.
A/RES/54/244 [➔ United Nations. 1999a] of 23 December 1999 entitled Review of the implementation of General Assembly resolution 48/218 B	Reaffirms OIOS's internal oversight responsibilities.
ST/SGB/2000/8 [➔ United Nations. 2000a] of 19 April 2000 entitled UN Regulations and Rules Governing Programme Planning, the Programme Aspects of the Budget, the Monitoring of Implementation and the Methods of Evaluation (PPBME)	Mandates OIOS-IED to submit to the GA, through the CPC, biennial reports on "strengthening the role of evaluation and the application of evaluation findings on programme design, delivery and policy directives." Further regulates the OIOS-IED evaluation function and provides a first glossary of terms.
ST/SGB/2002/7 [➔ United Nations. 2002a] of 16 May 2002 entitled Organization of the Office of Internal Oversight Services	Restates the organizational structure of OIOS, establishing a Monitoring, Evaluation and Consulting Division (MECD) within the Office.
A/RES/59/272 [➔ United Nations. 2004a] of 23 December 2004 entitled Review of the implementation of General Assembly resolutions 48/218 B and 54/244	Establishes that original versions of OIOS reports not submitted to the GA are available to any Member State upon request. Further provides that, when access to a report would be inappropriate for reasons of confidentiality or the risk of violating the due process rights of individuals involved in administrative proceedings within the Organization, the report may be modified, or in extraordinary circumstances withheld, at the discretion of the OIOS USG, who will provide reasons for this to the requesting party.
A/RES/61/245 [➔ United Nations. 2007a] of 7 March 2007 entitled Comprehensive review of governance and oversight within the United Nations and its funds, Programmes and specialized agencies	Formally creates the United Nations Inspection and Evaluation Division (IED) as a successor to the former Monitoring, Evaluation and Consulting Division (MECD).
A/RES/61/275 [➔ United Nations. 2007b] of 31 August 2007 entitled Terms of reference for the Independent Audit Advisory Committee and strengthening the Office of Internal Oversight Services	Establishes the Independent Audit Advisory Committee (IAAC) of the United Nations to act as a subsidiary body of the GA, to serve in an expert advisory capacity, and to assist the GA in fulfilling its oversight responsibilities. Mandates the IAAC to examine OIOS's work plan and budget proposal, and to advise the GA on the effectiveness, efficiency and impact of audit and other oversight functions.
A/64/263 [➔ United Nations. 2010a] of 29 March 2010 entitled Review of the implementation of General Assembly resolutions 48/218 B, 54/244 and 59/272	Requires OIOS to present the full text of comments received on all GA draft reports. Also required OIOS to define and compile a list of key oversight terms used in OIOS activities [IED #3].
A/RES/68/21 [➔ United Nations. 2013a] of 11 December 2013 entitled Report on the activities of the Office of Internal Oversight Services	Records the GA as taking note of the OIOS annual report in which the key oversight terms were reported.

1.2   The OIOS-IED Programme Impact Pathway (PIP)

The central document guiding OIOS-IED in implementing its mandate is the OIOS Strategic Framework (Part I, Section 1.1). Approved by the GA on a biennial basis, the Strategic Framework is the official document Member States use to set expectations on the overall results they want OIOS to achieve, along with the indicators they want OIOS to use in reporting on its performance against these targeted results. It also includes a brief summary of OIOS's overall strategies for achieving these results.

Although a vital document, the Strategic Framework is only a starting point. With a narrative devoted to OIOS-IED of only 315 words, the document's brevity alone limits its utility as a tool that OIOS or IED can use to optimally manage themselves toward the results Member States expect to see, or to monitor their own performance and correct course throughout the biennium. Particularly given the complexities OIOS faces within the UN environment in which it operates, what is needed is a tool that spells out its ultimate various objectives and how they relate to each other, how it intends to achieve these objectives in light of the challenging dynamics of the UN environment, and how it will know whether or not it is achieving these objectives.

In order to fill this gap, in 2012-13 OIOS developed Programme Impact Pathways (PIPs) for each of its three divisions, and the measures each division will use to monitor its own progress so as to learn and improve along the way. Figure 2 illustrates the OIOS-IED PIP.

The PIP is a "results roadmap" of sorts. Its elements can be viewed as follows:

- OIOS-IED's targeted impacts are the "reason for taking the trip" – that is, the ultimate change in the state of affairs OIOS-IED seeks to contribute to through its work;

- The outcomes OIOS-IED seeks to achieve toward these impacts are the "destination" on the roadmap and the "stops on the way" – i.e., the more immediate changes it aims to directly influence through its work short of the ultimate aim;

- The activities OIOS-IED undertakes and the resulting outputs it produces to achieve its targeted outcomes are the "road" and "directions" to take to get to the destination – i.e., the things its does, and the concrete things that result from doing them, through which it aims to exert change;

- The inputs at its disposal for undertaking the activities and producing the outputs on time, on budget and on target are the "driver's driving skills," "gasoline" and other key elements of a successful journey; and

- The assumptions it makes and drivers it envisages being at play in its drive for results are the "road conditions," "weather forecast," "signage" and other factors it foresees as helping or hindering it in arriving at its destination in a timely and "fuel-efficient" way.

In this scenario, the indicators OIOS-IED uses to measure its progress along the way can be viewed as the "vital statistics" of the trip - i.e., the regular check-ins on the GPS or map in order to monitor its location against where it should be given the time on the clock, the monitoring of how full the gas tank is or how awake the driver is, and so on.

In short, the PIP underscores OIOS-IED's reason for being. The Division exists not merely to produce evaluations, or even to produce high-quality evaluations, but rather to produce evaluations that are timely, objective, credible and relevant, and therefore used to inform decision-making on the part of its key stakeholders (Member States, the S-G and UN Secretariat programme managers) in ways that improve these programmes' relevance, effectiveness, impact and efficiency (evaluation criteria). In short, it aims to produce high-quality evaluations that make a difference.

Figure 2: IED Programme Impact Pathway (PIP), 2014-2015
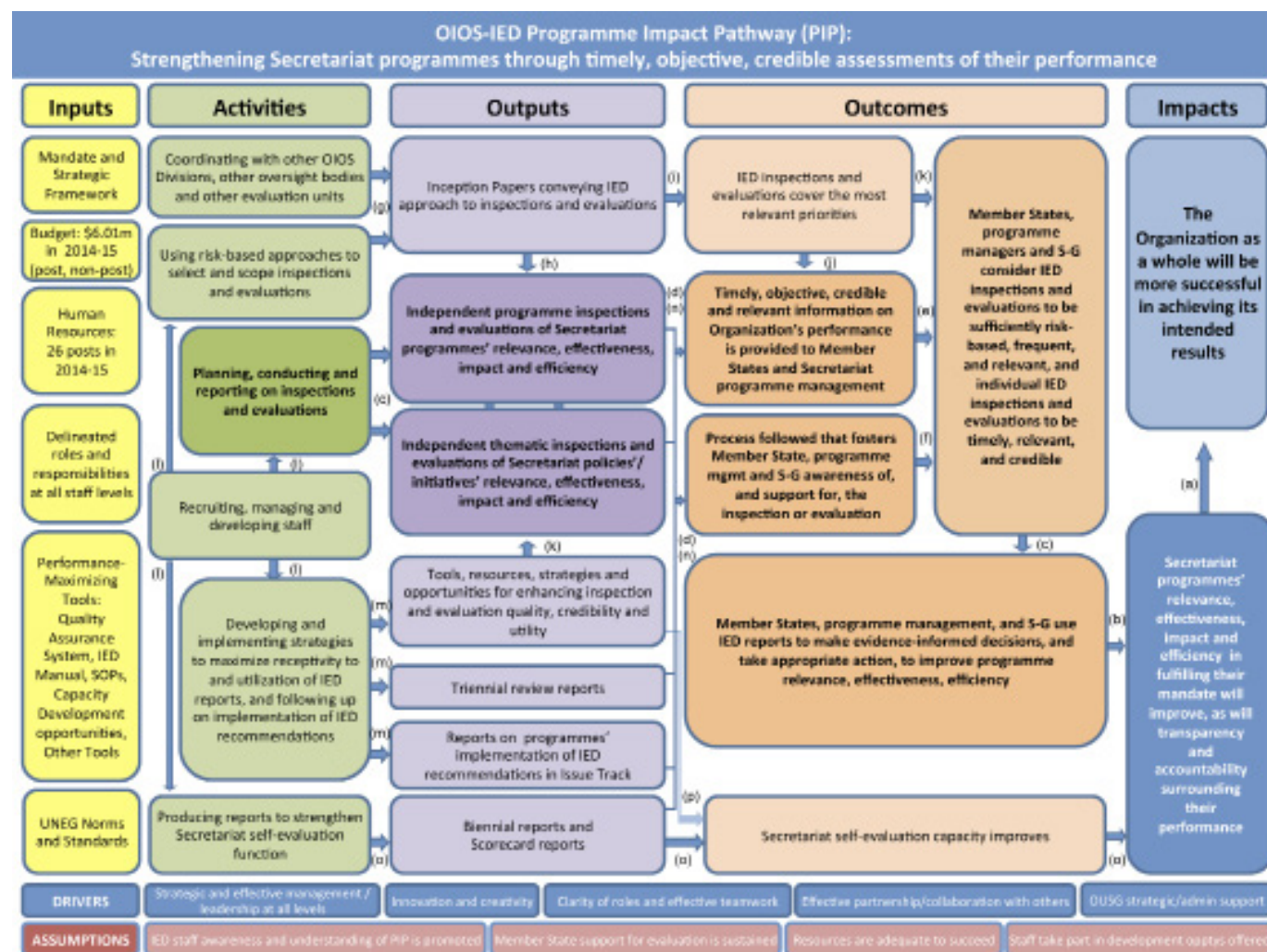
When OIOS-IED speaks of its aim for timeliness, objectivity, credibility and relevance in its evaluations, it means the following:

- Timeliness - Evaluations meet their pre-determined deadlines with no slippage, and the information they contain is conveyed to key stakeholders at the optimal moments for influencing key decisions;

- Objectivity - Evaluation results, conclusions and recommendations drawn are based solely on a logical analysis of the best evidence at hand, without undue influence of key stakeholders on the evaluators (independence) or of the evaluators' own biases (impartiality);

- Credibility - The products OIOS-IED generates and the processes it follows (i.e., for engaging with key stakeholders throughout the evaluation so that its products are ultimately used) are of a high quality both technically and strategically, ensuring that stakeholders will deem the analysis believable, and therefore worth accepting and acting on (or, alternatively, not easily ignored or disputed); and

- Relevance - Subjects selected for evaluation are those UN Secretariat programmes and themes of highest priority and highest risk to the Organization. Individual evaluations are scoped to inform the highest internal risks facing the evaluand and to inform the most critical considerations and decisions at hand.

Figure 3 provides a graphic overview of where OIOS-IED aims for all of its evaluations to be - i.e., in the bottom-right quadrant where evaluations are timely, objective, credible and relevant, and therefore used to inform decisions that strengthen UN Secretariat programmes. Each cell in the quadrant describes the typical state of affairs in each of four scenarios.

Figure 3:
OIOS-IED's Targeted Result: High-Quality, High-Utilization Evaluations That Make a Difference

Evaluation Utilization by Key Stakeholders

	Low	High
Evaluation Timeliness, Objectivity, Credibility and Relevance — Low	Poor product and/or process result in low utilization, with time and resources wasted (both for OIOS-IED and evaluand) that could have been spent elsewhere. In addition, the reputation of OIOS-IED and the evaluation function is compromised.	Despite a poor product and/or process, utilization is high, with the resulting decisions (some of which affect people's lives) resting on weak evidence. In addition, the notion of what constitutes credible evaluation is compromised.
High	Despite an excellent product and process, utilization is low. This might be despite OIOS-IED's best efforts to engage stakeholders. Although there is little risk to OIOS-IED's or the evaluation function's reputation, time and resources have been wasted.	Excellent product and process lead to high utilization, with important decisions resting on strong evidence. The reputation of OIOS-IED and of the evaluation function is enhanced.

Various sections of this manual provide guidance to help OIOS-IED ensure that its products and processes result in evaluations that are situated in the bottom-right quadrant of Figure 3. More concretely, Part II applies these concepts and principles to OIOS-IED's work from a practical standpoint. Walking the user through the full lifespan of an OIOS-IED evaluation, it describes the procedures OIOS-IED follows, step by step, to ensure that its evaluations are consistently timely, objective, credible and relevant – and used.

OIOS-IED has developed an in-depth narrative to accompany its PIP [IED #4], explaining the various causal pathways, along with indicators for measuring OIOS-IED's success. Coupled with the OIOS Strategic Framework, the OIOS-IED PIP and its indicators serve as critical management tools to help OIOS-IED steer itself toward the achievement of the results expected of it in the Strategic Framework. Accordingly, it is critical that all OIOS-IED staff understand and use these documents in their day-to-day work. Toward this end, beginning in early 2014, and based on a process of participatory consultation, OIOS-IED began using the PIP, along with the performance indicators associated with it, as a day-to-day management tool to help sharpen its focus on results and strengthen its drive toward results.

The OIOS-IED PIP also serves an important accountability function. Since 2014, OIOS-IED reports against its PIP and associated indicators to the Office of the OIOS Under-Secretary-General (OUSG) and the Independent Audit Advisory Committee (IAAC). Since 2008, OIOS also has a Self-Evaluation Policy [IED #5] in place that commits the Office and its divisions to subject their work to periodic external evaluation. The OIOS-IED PIP and the data OIOS-IED collects against the PIP indicators are the cornerstones of such evaluation efforts.

1.3   OIOS-IED's Oversight Universe
OIOS-IED undertakes independent inspections and evaluations (Part I, Section 1.1) of all UN Secretariat activities, covering 38 UN Secretariat departments and programmes

as well as peacekeeping operations (PKOs) and special political missions (SPMs).

Like the other OIOS divisions, OIOS-IED's mandate does not include UN funds, programmes or specialized agencies. However, as the OIOS-IED PIP (Part I, Section 1.2) indicates, this does not mean that OIOS-IED does not coordinate or communicate with their representatives. It does so in a variety of ways - e.g., informing them of its work and including them as stakeholders in individual evaluations. Moreover, to the extent that OIOS-IED finds evidence that aspects of the funds and programmes are key factors influencing the work of Secretariat programmes it evaluates, OIOS-IED can cite them as such in its reports.

In UN parlance, "Secretariat programme" is synonymous with "Secretariat department" or related structural entity, and not the term typically used to denote an initiative or cohesive set of project interventions. Within each programme are found a series of subprogrammes, which in turn denote the department's divisions, units, or similar substructures.

The work undertaken by the Secretariat programmes, which OIOS-IED inspects or evaluates, can be categorized into four main types. These are:

- Normative work;
- Analytical work;
- Operational work; and
- Internal support services.

Every UN Secretariat programme typically undertakes one or more of these types of work. Table 2 provides an overview of each of these types of work. It includes the typical outputs that programmes produce within each type and the corresponding outcomes that might be expected, and hence the types of results chains OIOS-IED would typically seek to assess in its evaluations. Table 2 also includes examples of programmes falling into each category. They are

included as examples only, and do not constitute a comprehensive listing of all UN Secretariat programmes falling into a given category.

of programmatic learning. Yet others fall somewhere in between these two models. How well different types of self-evaluation units are positioned to exercise their role in contributing to strengthening performance within their respective programmes is the subject of OIOS-IED's Biennial Reports and Scorecard Reports (Part I, Section 1.1).

Table 2: Typology of Secretariat Programmes' Areas of Work

Type of Work		Typical Outputs	Typical Outcomes	Typical UN Entity
Normative	Global	Global summits, international laws, standards and treaties	Consensus statements, ratification of conventions	OLA, UNCTAD, OHCHR, UNEP, UN-HABITAT
	Situational	Peace negotiations	Ceasefire agreements	DPA, SPMs, DPKO
Analytical		Publication of reports, statistics	Issue awareness, change in national policies/legislation	DESA, UN Regional Commissions
Operational	Peacekeeping	Military and police patrols	Reduction in violence and hostilities	PKOs (supported by DPKO/DFS)
	Humanitarian	Emergency response, coordination of the inter-agency response	Life-sustaining interventions, safe living conditions (health, mortality)	OCHA, OHCHR, UNHCR, UN-HABITAT
	Capacity-building	Diagnosis, training and advice to national authorities	Improved capacity of national institutions	UNODC, UNEP
Internal Support Services		Completion of UN recruitment; bookkeeping, travel arrangements; facilities and services	Efficient UN operations	DM, DGACM, OHRM

Some of these Secretariat programmes possess their own internal self-evaluation units. Such self-evaluation units are structurally diverse. Some are situated within their respective programmes' internal oversight function and report directly to its governing body, and thus exercise a level of operational independence that enables them to focus on strengthening accountability as well as learning. Others are embedded within and report to their respective programmes' management structures, and thus are more beholden to what management wants them to evaluate for the purpose

OIOS-IED's evaluations differ from those conducted by the self-evaluation units, in that OIOS-IED is operationally independent of the individual programmes it evaluates. OIOS's USG is appointed by the S-G following consultations with Member States, and this appointment is presented to the GA for approval. OIOS-IED evaluation reports are typically presented to an intergovernmental body - generally to OIOS's governing body, the GA Fifth Committee, often through the UN Committee for Programme and Coordi-

nation (CPC). While OIOS-IED, as per its mandate, is primarily focused on accountability it actively seeks to foster learning throughout its evaluations so that UN Secretariat programme managers are actively involved and therefore more receptive to - and more likely to use - OIOS-IED's evaluation recommendations.

Compared with self-evaluation units as a whole, OIOS-IED brings a broader evaluation scope focused on the performance of the programme as a whole, rather than on specific subprogrammes or projects within it. Compared to those self-evaluation units with little or no operational independence, OIOS-IED is also able to ask and answer evaluation questions that need to be addressed, rather than those that programme management consents to being addressed, and is able to report its conclusions and recommendations as it sees fit.

Figure 4 delineates OIOS's scope and independence as compared to the Joint Inspection Unit and programme-specific self-evaluation units.

Figure 4: Scope and Independence of UN System Evaluation Bodies
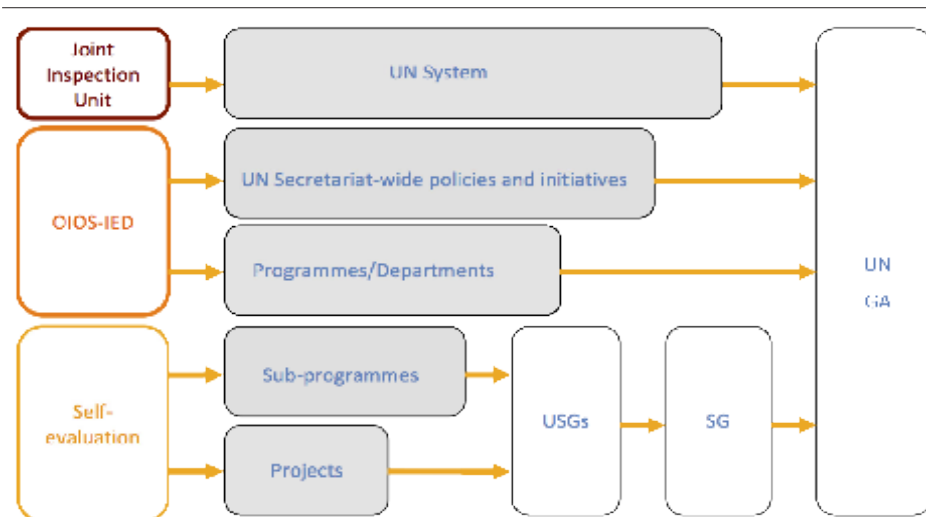
1.4   OIOS-IED Products

OIOS-IED's main products are inspections and evaluations. But what are these exactly, what subtypes of these exercises are there, and how do they differ from each other? What other products does OIOS-IED produce and toward what ends?

This section describes the full range of outputs OIOS-IED produces in the interest of generating timely, objective, credible and relevant information on UN Secretariat programmes' performance that Member States, the S-G and programme management use to improve such performance.

OIOS-IED's main products are as follows:

- Programme evaluations;
- Thematic evaluations;
- Inspections;
- Ad hoc inspections and evaluations;
- Biennial Reports on the State of Evaluation in the UN Secretariat;
- Evaluation Scorecards;
- Triennial Reviews; and
- Other products.

PROGRAMME EVALUATIONS

Full programme evaluations, also referred to as "in-depth" evaluations when mandated by the CPC, assess the overall relevance, effectiveness, impact and efficiency of a single Secretariat programme, subprogramme, or of a peacekeeping operation (PKO) or special political mission (SPM). The CPC expects OIOS-IED to have inspected or evaluated the full universe of 38 Secretariat programmes every eight years, an expectation that helps shape the formulation of OIOS-IED's risk-based work programmes (Part I, Section 1.2).

THEMATIC EVALUATIONS

Thematic evaluations typically assess a cross-cutting theme or activity (e.g., implementation of a gender mainstreaming policy or knowledge management) across multiple Secretariat programmes or PKOs/SPMs. They can also assess the cumulative effects of multiple programmes sharing common objectives and purposes (e.g., the Secretariat's contribution to the Millennium Development Goals) or the effectiveness of coordination and cooperation among different programmes (e.g., the interaction of UNHCR, OCHA and Secretariat members of the Emergency Shelter and Protection Clusters in humanitarian action).

Together, programme and thematic evaluations make up the majority of OIOS-IED's reports, and the key products driving its main results pathway on the OIOS-IED PIP (Part I, Section 1.2).

INSPECTIONS

Inspections are shorter, more focused and more targeted reviews of an organizational unit, issue or practice perceived to be of potential risk, in order to determine the extent to which it adheres to established norms, good practices or other pre-determined criteria, and to identify corrective action as needed. In practice, OIOS-IED's inspections and evaluations are similar in nature, differing mainly in their scope and duration. In addition, inspections tend to be less regularly planned.

OIOS-IED inspections are not the same as investigations, which focus on determining wrongdoing within the Organization, and which require referral to the competent authorities. Nor are they physical inspections for the monitoring and verification of compliance with international obligations, such as inspections for weapons of mass destruction.

AD HOC INSPECTIONS AND EVALUATIONS

Ad hoc requests for inspections or evaluations are made by any of the Organization's stakeholders, including Member States, the S-G and Secretariat programme managers. They are undertaken subject to OIOS-IED's review of the proposed topic's strategic importance and potential risk to the Organization, and consideration of the resource implications of fulfilling the request.

In most cases, ad hoc evaluations are identical to standard programme or thematic evaluations, and differ only in the demand-driven way in which they come about. In some cases, however, they differ slightly from standard evaluations in that they might focus on a specific section or unit of a subprogramme, a programme-specific policy, or some other level of analysis than either a programme or a theme. Often OIOS-IED is asked to conduct ad hoc inspections or evaluations in order to bring a greater degree of operational independence than the concerned self-evaluation unit possesses.

BIENNIAL REPORTS ON THE STATE OF EVALUATION IN THE UN SECRETARIAT

In accordance with ST/SGB/2000/8 Table 1, OIOS-IED is mandated to submit to the GA, through the CPC, Biennial Reports on "strengthening the role of evaluation and the application of evaluation findings on programme design, delivery and policy directives." Together with the Evaluation Scorecards, the Biennial Reports are thus the main products through which OIOS-IED seeks to strengthen self-evaluation capacity within the UN Secretariat, a mandated aspect of its work enshrined in the OIOS Strategic Framework (Part I, Section 1.1) and depicted in its PIP (Part I, Section 1.4).

Biennial Reports prior to 2008 focused on reviewing both internal programme self-evaluation and central evaluation practice and capacity in the Secretariat. Since 2008, Biennial Reports also provide a synthesis of the results of all Secretariat programme self-evaluations. They typically include the following sections:

- An assessment of the current capacity, quality and utility of the evaluation function within the Secretariat;
- A meta-analysis of key results, conclusions and recommendations from evaluation reports finalized in the biennium covered; and
- A presentation of the OIOS-IED work plan for the coming biennium.

Box 1 summarizes the basic steps involved in undertaking a review leading to a Biennial Report.

Box 1: Basic Steps for Conducting a Biennial Review

1. Request evaluation reports from all Secretariat entities for the biennium in question (providing a definition of what constitutes an evaluation);

2. Screen the reports [IED #6] to determine which documents are evaluation reports and which are not;

3. For those reports passing the screening, conduct an assessment of their quality [IED #7] based on the UN Evaluation Group (UNEG) norms and standards (Part I, Section 1.5). Consider pulling a purposive sample (Part II) to reduce the number of reports assessed, based on time and resources;

4. For those reports deemed to be of "excellent" or "good" quality, conduct an analysis of their contents within the context of the key strategic priorities of the Organization;

5. Conduct a survey of all evaluand focal points for the biennial review to provide quantitative data on evaluation capacity and practice;

6. Conduct interviews with all evaluand focal points to provide qualitative data on evaluation capacity and practice;

7. Conduct an assessment [IED #8] of the entities' evaluation policies; and

8. Review budget fascicles to obtain data on evaluation.

EVALUATION SCORECARDS

In 2013, OIOS-IED introduced a new product for all Secretariat entities - Evaluation Scorecards [IED #9]. The Scorecards provide a programme-by-programme assessment of evaluation capacity and practices, based on 15 indicators emanating from the UNEG norms and standards (Part I, Section 1.1). Together with the Biennial Reports, the Evaluation Scorecards are the main products through which OIOS-IED seeks to strengthen self-evaluation capacity within the UN Secretariat, a mandated aspect of its work enshrined in the OIOS Strategic Framework (Part I, Section 1.1) and depicted in its PIP (Part I, Section 1.2).

TRIENNIAL REVIEWS

Triennial Reviews are follow-up exercises, undertaken three years after every GA-mandated inspection and evaluation report, in accordance with a decision by the CPC at its 22nd session to review the implementation of its recommendations. These reviews involve the collection of evidence to verify implementation of recommendations and to describe how recommendations have been implemented. A Triennial Review is usually started in December and completed in March of the following year in order to be presented to the CPC in June. Peacekeeping reports are not generally mandated by the GA and so are not subject to Triennial Reviews.

Box 2 summarizes the basic steps involved in undertaking a Triennial Review.

Box 2: Basic Steps for Conducting a Triennial Review

1. Review CPC report that endorsed the evaluation recommendations to determine whether any of these were altered (the CPC may add its own recommendations, or change the substance of an existing recommendation);

2. Print out and review all Issue Track entries since the report was issued;

3. Meet with main report author to obtain an accurate understanding of the intent and substance of the recommendations;

4. Develop a matrix to outline, by recommendation, the follow-up action and evidence that is required to verify implementation of each recommendation;

5. Obtain evidence using interviews, document or website reviews, and/or surveys;

6. Collect evidence and make final conclusion on status of implementation for each recommendation; and

7. For recommendations not implemented, assess the reason(s) and implications of non-implementation. For implemented recommendations, assess the impact of implementation (if feasible).

OTHER OIOS-IED PRODUCTS

OIOS-IED produces a number of other products. As its PIP (Part I, Section 1.2) conveys, most of these are intended either to strengthen the timeliness, objectivity, credibility and relevance of the Division's main products or to foster an environment conducive to OIOS-IED's work. In this way, other OIOS-IED products, such as this Inspection and Evaluation Manual and the numerous sources of OIOS-IED-designed guidance appended to it, also constitute important products that help it achieve its targeted results.

Another major output OIOS-IED produces is the inception paper (Part II, Section 1.1) that precede almost every programme and thematic evaluation. Inception papers define the scope of the broad evaluation topic, along with the associated methods, overall approach and specific practical considerations facing the evaluation. In the interest of transparency, they also convey OIOS-IED's rationale behind its decisions in defining the approach to the evaluation.

Owing to their frequently shorter cycle, OIOS-IED inspections and some evaluations requested by UN Secretariat programme managers (e.g., some peacekeeping evaluations) may be preceded by a shorter scoping and planning document - i.e., a terms of reference, or ToR - rather than the longer inception paper.

1.5   UNEG Norms and Standards

The UN Evaluation Group (UNEG) is a professional inter-agency network that brings together the evaluation units of the UN system, including specialized agencies, funds and programmes, and affiliated organizations. In 2013, UNEG had 43 members and three observers. OIOS-IED has been a member of UNEG since its inception, with many of its management and staff serving as chairs or members of task forces on various topics such as evaluation practice exchange, norms and standards, evaluation of normative work, and human rights and gender equality. OIOS-IED's Director is serving as UNEG Chair from 2012 though 2015, during which she oversaw the formulation of a 2014-2019 Strategy for UNEG [➔ UNEG. 2013a].

In April 2005, the UN endorsed the UNEG norms and standards for evaluation in the UN system. As its PIP (Part I, Section 1.2) indicates, OIOS-IED relies on the UNEG norms and standards as important inputs into its work that help guide its inspections as well as its evaluations (Part I, Section 1.4).

UNEG NORMS
The UNEG norms [➔ UNEG. 2005a] seek to ensure that evaluation entities within the UN follow agreed-upon basic principles. They provide a reference for strengthening, professionalizing and improving the quality of evaluation in all entities of the UN system. Figure 5 summarizes the 13 norms for evaluation in the UN system.

Figure 5: Norms for Evaluation in the UN System

Although all of the norms are crucial, three of these warrant particular attention: independence, impartiality, and transparency.

Independence

For OIOS-IED, its independence is critical to the successful exercise of its mandate. In addition to operational independence, OIOS-IED aspires to achieve the highest level of behavioural independence possible. Compared to operational independence, which has to do with the way OIOS-IED is institutionally structured and mandated to enable it to initiate, carry out and report on its work, behavioural independence has to do with how OIOS-IED's staff are individually enabled to conduct their work throughout the evaluation without undue interference by those involved in implementing the programme, project, policy or other unit of analysis being evaluated.

Even with a mandate for operational independence, OIOS-IED staff sometimes face

situations where various stakeholders attempt to influence evaluations in one way or another - i.e., beyond the appropriate ways of doing so, such as providing their insights, information and materials during data collection.

In short, behavioural independence ensures that OIOS-IED's analysis flows solely from the best possible evidence at hand, rather than a skewed or truncated subset of such evidence owing to stakeholder attempts to unduly influence the evaluation. It is thus a vital element in ensuring the evaluation's credibility and objectivity.

Table 3 provides an overview of specific "do's and don'ts" to help OIOS-IED evaluation teams maximize behavioural independence during an evaluation.

Table 3: Do's and Don'ts for Ensuring Behavioural Independence

Do ...	Do NOT ...
• Ensure that OIOS-IED's operational independence is clearly stipulated in the evaluation inception paper or ToR, as per the OIOS-IED Quality Assurance Checklist	• Succumb to undue pressure from any stakeholders, during the consultation process, to change any aspects of the inception paper or ToR without a rationale the OIOS-IED evaluation team deems to be well-founded
• Liaise with evaluand focal points and other stakeholders to convey the independence of the evaluation and what this means	• Allow veto power over any aspect of the evaluation, while consulting with stakeholders for their inputs on the evaluation,
• Monitor other stakeholders' adherence to their responsibilities for independence, and ensure appropriate corrective action	• Wait until the end of the evaluation to flag threats to independence
• Communicate unresolved stakeholder infringements on the evaluation's independence to the Team Leader and Section Chief, in order to determine appropriate action (including, in extreme cases, mentioning non-cooperation in the evaluation report)	• Share data collection instruments with any stakeholders until data collection is complete, except in the pre-testing of these instruments, in order to reduce opportunities for information leaks

Of these, the most important to remember is that OIOS-IED staff must alert their re-spective Chief of Section (also referred to as "Section Chief") when threats to behav-ioural independence occur, in order to identify the best strategies for managing them.

It is important to underline that independence, whether operational or behav-ioural, is not the same as isolation or autonomy from the evaluand. In fact, the opposite is true - independence and consultation are equally vital complemen-tary elements of OIOS-IED's strategy for achieving the results outlined in its PIP. Without independence, OIOS-IED cannot achieve objectivity. Without consulta-tion, it cannot achieve relevance. Without either, it cannot achieve credibility.

Impartiality

Impartiality is another of the UNEG norms of particular importance to OIOS-IED. Like independence, impartiality plays a central role in strengthening the credibility and objectivity of OIOS-IED's evaluations. Impartiality is distinct from independence. Whereas independence has to do with "insulating" the evaluation from undue external influence, impartiality has to do with monitoring and addressing the evaluator's own potential bias in the evaluation.

Impartiality is the absence of evaluator bias in methodology, in the consideration and presentation of the evaluand's achievements and challenges, and in the reflection of different stakeholder views and pertinent evidence in the evaluation report. OIOS-IED staff are recruited, managed and developed to maintain an im-partial approach to their work. However, bias can never be fully eliminated. Being aware of, and advancing, impartiality helps strengthen the evaluation's credibil-ity and objectivity, and therefore its utilization, by limiting the likelihood that the evaluator's own personal preferences or interests unduly influence the analysis.

Conflicts of interest are one of the most common sources of bias, and have the potential to be at play when any of the following scenarios is present:

- The evaluation team member has previously worked for the evaluand, or for one of its staff members involved in the programme under evaluation, and might therefore have (or be perceived to have) strong positive or negative pre-conceptions about one or the other;

- An evaluation team member has applied for a job with the evaluand before or during the evaluation, thus increasing the likelihood that s/he will look favourably on the programme (or be perceived to look favourably on it);

- A close family member works for the evaluand, thus raising similar questions as in the instances described above;

- Based on previous exposure to the evaluand or one of its staff members in-volved in the programme under evaluation, an evaluation team member has strong personal feelings (positive or negative) about either that will get in the way of his or her impartiality; and

- Other previous links to the evaluand that will compromise the evaluation team member's ability to be impartial.

Conflicts of interest can be either real or perceived, and even if one of the fore-going scenarios occurs, this does not automatically mean the evaluator cannot act with impartiality. What is most critical is that, as with threats to independ-ence, if staff members believe a conflict of interest (either real or perceived) might exist, they immediately notify their supervisor, who takes appropriate action to ensure that the evaluation process is not impaired. Any instances of undue influence on OIOS-IED evaluators to conduct their evaluative work in an impartial manner are reported to the OIOS USG for his or her action.

Transparency and consultation with evaluands

As the previous sections have suggested, OIOS-IED, though mandated with operational and behavioural independence, also takes a utilization-focused approach [➜ Quinn Patton. 2008a] [📖 BetterEvaluation] to its evaluations. OIOS-IED does not view its independence and its strong utilization focus as mutually exclusive. On the contrary, balancing its independence with a commitment to consultation is critical to ensuring that its evaluations are relevant and credible and therefore used.

OIOS-IED therefore seeks consultation and collaboration with evaluands, right from the outset of the evaluation planning stage and extending throughout the evaluation process. In addition, in its inception papers (Part II, Section 1.1), OIOS-IED's document for defining the scope of its evaluations, the Division transparently communicates the rationale of its main strategic choices for an evaluation. This transparency helps promote evaluator impartiality and further strengthens OIOS-IED's credibility with the evaluand.

The evaluand focal point role is a particularly important aspect of the OIOS-IED work process. Evaluand focal points are assigned by their respective programmes to fill the key liaison role on the evaluation at hand. They provide assistance throughout the evaluation process with:

- Organizing key meetings such as an entry and exit conferences;
- Providing key documents and other requested materials;
- Compiling staff lists;
- Providing lists of key stakeholders;
- Establishing internal reference groups, where applicable, and facilitating communication with them;
- Facilitating interviews with programme management and staff;
- Organizing field missions;
- Advising OIOS-IED on strategies to maximize the evaluation's utilization;

- Following up with any requests OIOS-IED has of the programme and ensuring timely compliance;
- Compiling the evaluand's comments on draft reports prior to their finalization by OIOS-IED; and
- Understanding their role in ensuring OIOS-IED's behavioural independence in the evaluation.

UNEG STANDARDS

The UNEG standards [➜ UNEG. 2005b] build upon the norms. They are drawn from UNEG member good practices, and are intended to guide the establishment of the institutional framework, the management of the evaluation function, and the conduct and use of evaluations. In summary, there are 50 standards for evaluation in the UN system. They fall within four broad categories:

- Institutional framework and management of the evaluation function;
- Competencies and ethics;
- Conducting evaluations; and
- Reporting.

Anonymity and confidentiality

In line with UNEG Standard 2.7, OIOS-IED staff members protect the anonymity and confidentiality of the information they receive. Anonymity refers to the protection of an individual's identity in the evaluation, such that their individual identity is never known - even to the evaluators. This is typically achieved only in large-sample surveys in which respondents are not known to the evaluation team. Confidentiality, in contrast, refers to the safeguarding of individuals' identities, to respecting individuals' right to provide information in confidence, and to ensuring that sensitive information cannot be traced to its source.

Anonymity and confidentiality are important concepts from both a methodological and an ethical standpoint. Methodologically speaking, evaluators obtain less valid and

reliable information from interviewees or respondents who fear their feedback might be shared. Validity refers to the accuracy of an assessment - whether or not it measures what it is supposed to measure. Reliability, by contrast, refers to the notion that data collected are consistent - from one time period to the next, or from one interviewee or respondent to the next. What is being asked and what is being given in reply does not change in interpretation. (Part II, Section 4.2)

In OIOS-IED's work, this can sometimes be a challenge - e.g., when interviewing women in cultures where meeting alone, particularly with a male interviewer, is considered unacceptable. Ethically speaking, the failure to safeguard individuals' identities can have serious negative consequences for those individuals, particularly in already-vulnerable populations and particularly when dealing with sensitive subject matter.

OIOS-IED employs a number of measures to protect confidentiality during data collection and analysis. These include, but are not limited to, the following:

- Informing individuals up front of OIOS-IED's policy of confidentiality, and of any foreseen limits to their confidentiality;

- Keeping interview and focus group notes containing any identifying information on a computer's hard drive, external drive, or a password-protected folder on OIOS-IED's server;

- Refraining from sharing specific information provided during an interview or focus group, in ways that would reveal an individual's identity, with anyone outside the OIOS-IED evaluation team and especially with those who might use such information to harm the individual in any way;

- Knowing when to employ individual interviews as opposed to focus groups;

- Not allowing programme staff to attend the interviews of other interviewees unless OIOS-IED deems such attendance serves a specific purpose for the evaluation; and

- Not citing individuals in a report in a way that might reveal their identity - i.e., not only by mentioning them by name, but also by mentioning identifying characteristics or quoting them with jargon or syntax they are known to use.

1.6   Ensuring Human Rights and Gender-Sensitive Evaluations

In line with GA mandates and the UN Evaluation Group (UNEG) norms and standards (Part I, Section 1.5), OIOS-IED integrates human rights and gender perspectives into its evaluation practice. Human rights and gender equality-responsive evaluation has two essential elements. It is about what the evaluation examines and how the evaluation is undertaken. It is not an approach just for programmes with an explicit focus on human rights and/or gender equality, but rather provides a holistic and meaningful assessment of any and all programming.

OIOS-IED strives to integrate human rights and gender equality throughout the different steps of its evaluation process, including: putting together evaluation teams balanced on gender and other dimensions, wherever feasible; mapping evaluation stakeholders with a view to gender and human rights considerations; reflecting any relevant human rights or gender-specific aspects of the programme in the evaluation Programme Impact Pathway (PIP) or Thematic Impact Pathway (TIP) Part II; ensuring appropriate questions are included in the inception paper or ToR that focus on these dimensions, and framing them accordingly; and collecting disaggregated data and writing the evaluation report. In doing so, OIOS-IED intends to contribute to strengthened accountability for and learning within the Organization on what is and is not working in its quest to advance human rights and gender equality.

There are a number of resources on integrating human rights and gender equality in evaluation, including those produced by the UNEG, with OIOS-IED's involvement [➔ UNEG. 2011a] [➔ UN Women. 2010a] [➔ UNICEF. 2011a] [📖 Gender and Evaluation] [📖 Feminist Evaluation] [📖 My M&E] [📖 BetterEvaluation]. OIOS-IED staff members (re-)familiarize themselves with these important sources of guidance right at the outset of every new evaluation.

1.7   OIOS-IED Staff and Financial Resources

The foregoing sections describe what OIOS-IED does and what it produces - i.e., the activities and outputs in its PIP (Part II, Section 1.2) and to what ends - i.e., its targeted outcomes and impacts in the PIP. The present section speaks to a number of the key inputs it relies on to achieve its targeted results. OIOS-IED's ability to effectively harness human and financial resources is the linchpin of its drive to achieve results - and to achieve these results in the most effective, efficient and cost-effective ways possible.

HUMAN RESOURCES

Figure 6 depicts OIOS-IED's organigramme. As of 2014-2015, OIOS-IED is comprised of 26 staff (22 Professional and 4 General Service staff), embodying a wide range of professional backgrounds. OIOS-IED staff have experience conducting evaluations in a broad range of thematic areas (e.g., economic and social development, health, peacekeeping, humanitarian action, the environment), and have worked in a variety of fields including development, economic affairs, programme and project management, planning, monitoring, public policy and administration, law and communications.

OIOS-IED's Director bears ultimate responsibility for all inspections and evaluations (Part I, Section 1.4) in the Division. S/He reports directly to the OIOS USG, who in turn reports to the S-G and submits reports in his or her own name to the GA.

OIOS-IED's human resources are its most valuable asset in its quest for results.

It is therefore vital that OIOS-IED staff are well equipped with both the technical and strategic skills appropriate to their staff level to contribute to these results. OIOS-IED does so through strategic recruitment, staff induction and development processes, and through strong performance management, to ensure that the right staff are hired and are provided the training, supervision and guidance needed to succeed. In addition, OIOS-IED's management attempts to ensure that, all else held equal, OIOS-IED evaluation teams (Part I, Section 1.1) are staffed by those staff members who are best suited for the specific evaluation at hand.

Clarity of roles and responsibilities among its staff is a core input underpinning OIOS-IED's work. Toward this end, OIOS-IED relies on several useful documents UNEG has developed for bench-marking staff skills and professional development. These include core competencies for evaluators [➔ UNEG. 2008a] and core competencies for Heads of Evaluation in the UN system [➔ UNEG. 2008b] as well as generic job descriptions for evaluation staff at the P1-P2 [➔ UNEG. 2008c], P3 [➔ UNEG. 2008d], P4 [➔ UNEG. 2008e] and P5 [➔ UNEG. 2008f] levels.

In addition, OIOS-IED has developed two tools that help clarify roles and responsibilities within the Division. These include a matrix outlining the roles and responsibilities of the various staff categories [IED #10]. They also include a team compact (Part I, Section 1.1) to help OIOS-IED's diverse evaluation teams clarify specific roles and responsibilities, work styles, and professional development goals that each team member brings to the specific evaluation project at hand.

FINANCIAL RESOURCES

OIOS-IED's budget is drawn from both the Regular Budget (RB) of the UN and the peacekeeping support account (also known as the support account, or QSA, budget). Budgets are approved by the Fifth Committee, based on an OIOS submission approved by the Office of Programme Planning, Budget and Accounts (OPPBA) in the Department of Management (DM). For the 2014-2015 biennium, OIOS-IED received

USD 9.26 million in RB funds. Peacekeeping evaluation work is funded annually from the Peacekeeping Support Account. For the twelve months to June 2014, OIOS-IED received USD 996,900 in QSA funding.

2.   OIOS-IED's Approach to Work Planning

This chapter focuses on ways in which OIOS-IED seeks relevance and credibility in its evaluations, right from the start. First, it describes OIOS-IED's approach to overall work planning - how the Division determines the highest-priority, most relevant programmes and themes it will evaluate during the biennium to come. Second it builds on (Part I, Section 1.7), describing in greater detail how OIOS-IED allocates its staff, as well as consultants, to the evaluations in its work plan for maximum credibility.

This chapter has three main sections:

- OIOS-IED Strategic Risk Assessment Framework (Part I, Section 2.1)
- OIOS-IED Evaluation Teams (Part I, Section 2.2)
- Use of Consultants (Part I, Section 2.3)

2.1   OIOS-IED Strategic Risk Assessment Framework

The CPC expects OIOS-IED to inspect or evaluate the full universe of 38 Secretariat programmes every eight years. Within this overall aim for cyclical coverage, OIOS-IED takes a strategic risk-based approach in preparing its biennial work plans. This risk-based approach aims to ensure that the Division's evaluations are maximally relevant by addressing oversight and strategic priorities in a regular and timely manner, and by focusing finite esources on those areas requiring most urgent attention in the biennium at hand. It is the primary means by which OIOS-IED determines its biennial work plans. At the same time, evaluations are sometimes requested on an ad hoc basis by the S-G, by Member States or by UN Secretariat programme managers.

In coordination with other OIOS divisions and other oversight bodies as appropriate, OIOS-IED identifies and conducts a programme of evaluations based on
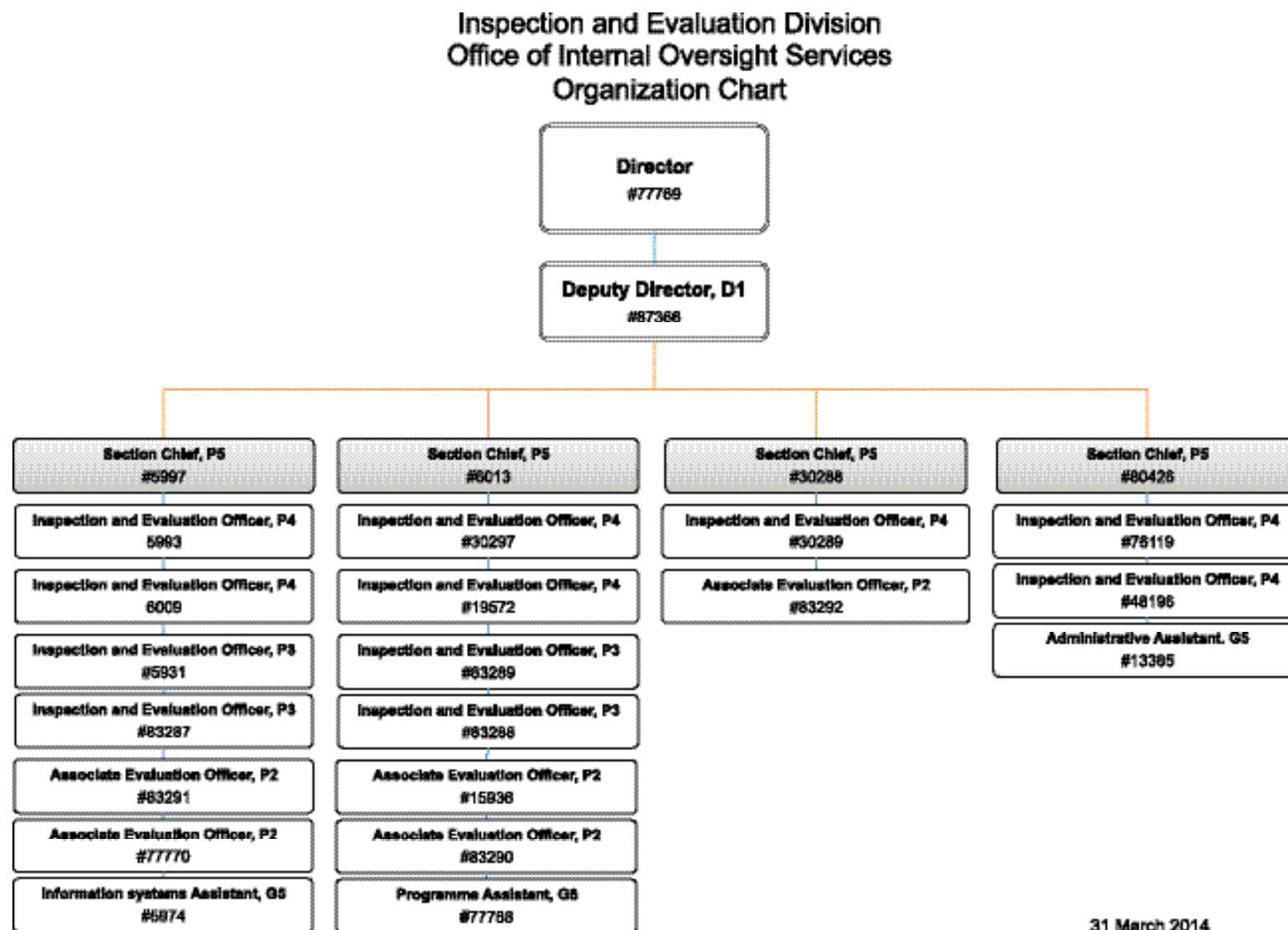
an assessment of the highest risks to and priorities of the Organization as well as a systematic and periodic coverage of Secretariat programmes. Owing to their separate budget stream and work cycle, OIOS-IED's peacekeeping-related evaluations are planned using a separate risk assessment. Starting in 2014, OIOS-IED is piloting a consolidated risk assessment exercise that encompasses both its regular (RB-funded) programme and thematic evaluations as well as its QSA-funded peacekeeping evaluations.

In selecting programme evaluations, OIOS-IED uses a planning framework that considers factors relating to three components:

- Risk that the programme will not be able to achieve its objectives and execute its strategies successfully;

- Priorities highlighted by the S-G, the Chief Executives Board (CEB) and GA; and

- Systematic and periodic coverage of Secretariat programmes to ensure sufficiently regular, independent and objective information on Secretariat programme results and the attainment of GA mandates, which is needed to support reflection and decision-making by the Organization's governance and management bodies.

OIOS-IED generates an annual risk assessment calculation, rank-ordering all Secretariat programmes by their overall risk, based on the following overarching criteria: i) the residual risk in seven sub-focus areas in governance, strategy and programme management; ii) programme budgets; and iii) monitoring and evaluation capacities. An identification of the Organization's priorities and a comparison with each programme's areas of activities results in a ranking of programmes by overall priority level.

Figure 6: The IED Organigramme, 2014-2015

Inspection and Evaluation Division
Office of Internal Oversight Services
Organization Chart

Director
#77769

Deputy Director, D1
#87366

Section Chief, P5 #5997	Section Chief, P5 #6013	Section Chief, P5 #30288	Section Chief, P5 #80426
Inspection and Evaluation Officer, P4 5963	Inspection and Evaluation Officer, P4 #30297	Inspection and Evaluation Officer, P4 #30289	Inspection and Evaluation Officer, P4 #76119
Inspection and Evaluation Officer, P4 6009	Inspection and Evaluation Officer, P4 #19572	Associate Evaluation Officer, P2 #83292	Inspection and Evaluation Officer, P4 #46196
Inspection and Evaluation Officer, P3 #5931	Inspection and Evaluation Officer, P3 #83289		Administrative Assistant, G5 #13365
Inspection and Evaluation Officer, P3 #83287	Inspection and Evaluation Officer, P3 #83288		
Associate Evaluation Officer, P2 #83291	Associate Evaluation Officer, P2 #15836		
Associate Evaluation Officer, P2 #77770	Associate Evaluation Officer, P2 #83290		
Information systems Assistant, G5 #5974	Programme Assistant, G6 #77768		

31 March 2014

An aggregation of both risk and priority rankings allows OIOS-IED to create a final ranking. Through this process, OIOS-IED categorizes programmes as being of high risk/priority, medium risk/priority and low risk/priority.

OIOS-IED may prescribe more frequent or targeted assessments of a given Secretariat programme in the event that its risk assessment yields specific risks, e.g., shortcomings in the programme's monitoring and evaluation capacity. OIOS-IED usually addresses these specific issues through smaller-scale inspections.

OIOS-IED selects topics for its thematic evaluations based on a systematic review and analysis of documents to identify the Organization's substantive and internal management priorities and upcoming issues. Topics for OIOS-IED's peacekeeping evaluation work are identified through consultations with the Department of Peacekeeping Operations (DPKO) and the Department of Field Support (DFS), an inception review of DPKO/DFS headquarters activities, and a thematic risk assessment.

OIOS-IED presents its biennial work programme to the CPC, from which the CPC selects those reports that it will later review. Reports the CPC will not review are called programme manager reports. They are directed to the evaluand and other stakeholders, as OIOS-IED deems appropriate. OIOS-IED submits both its RB- and QSA-funded work plans to the IAAC.

2.2   OIOS-IED Evaluation Teams

OIOS-IED takes a team-based approach in its work. The size and composition of each team assigned to a given evaluation is based on a consideration of the following factors, all of which are rooted in OIOS-IED's overarching concern of deploying the optimal team to achieve the results described in its PIP:

- The scope of the inspection or evaluation being conducted;
- The particular skill sets needed (such as, for example, language skills);
- Gender balance;
- Section resources and demands;
- Staff members' professional development aspirations as per their performance plan; and
- Staff availability to meet project deadlines.

Each evaluation team typically consists of:

- 1 Section Chief;
- 1 Team Leader;
- 1 or more team members, who might support more than one evaluation; and
- 1 Administrative Assistant, who supports more than one evaluation.

As indicated in OIOS-IED's organigramme (Part I, Section 1.7), Section Chiefs are P5-level managers. They are responsible for ensuring overall timeliness, relevance, objectivity and credibility of the inspections and evaluations in their portfolio. They manage, guide, support and directly assist the evaluation teams in their section, with a view to achieving results while ensuring staff accountability and fostering their development. Section Chiefs report directly to the OIOS-IED's Director, who is accountable for the quality and timeliness of all OIOS-IED reports submitted to the OIOS USG and OIOS's governing bodies.

Team Leaders are typically, but not always, P4 or P3-level inspection and evaluation officers. They have overall responsibility for successful completion of individual projects and manage their evaluation teams. Team Leaders report directly to Section Chiefs.

Team members are P3 and P2-level inspection and evaluation officers. They are responsible for assisting, in some cases working independently, on all stages of

the inspection or evaluation, including preliminary research, design, data collection and analysis, and report writing. Team members are supervised by Section Chiefs. They are guided and mentored by their Team Leader, who provides inputs into the Section Leader's performance appraisal of the Team Leader's team members.

Administrative Assistants support the evaluation teams in their respective sections, assisting with correspondence, travel, report formatting and processing, as well as with web-based surveys. They report directly to Section Chiefs.

Every staff member has a role to play in assuring quality, as per the OIOS-IED Quality Assurance System (QAS) (Part I, Chapter 3). An open process of communication is essential to ensure that all perspectives are shared and all ideas welcomed. If there is disagreement within the evaluation team on quality issues, the Section Chief is consulted, and if there is disagreement within the section, the Directorate (that is, OIOS-IED's Director and Deputy Director) is consulted. Although ultimate responsibility for quality rests with OIOS-IED's Director, all voices are valued.

At the outset of evaluations, evaluation teams develop a team compact [IED #11], the purpose of which is to:

- Share information about available resources for the evaluation;
- Encourage and facilitate team cohesion;
- Provide evaluation teams with a tool to better understand and utilize working styles of all team members; and
- Guide discussion among team members about their contributions and expectations about a given evaluation; Share information about available resources for the evaluations; and
- Conrtibute to the realization of individual team members' professional development goals.

2.3   Use of Consultants

Depending on the scope of an evaluation, OIOS-IED engages consultants in a variety of capacities, namely as:

- Subject-matter experts; and
- Advisory panel members.

Subject-matter experts are engaged for their substantive knowledge of specific technical aspects related to the evaluation - e.g., international humanitarian law, sanctions regimes or sampling strategies in post-conflict environments. They provide guidance and input into the project on a specific topical area, at key stages of the project, such as the inception stage, data collection and analysis, and/or draft report.

Project-specific advisory panels are established to include subject-matter and/or methodology experts. As the term suggests, advisory panel members provide input into evaluation projects at critical stages, such as during scoping, development of the inception paper, data collection and drafting of the evaluation report. The panels can be used in various capacities, depending on project needs. Some function as external reference groups of experts, whose role is to provide a non-UN perspective on the subject matter at hand. Others might include other evaluators familiar with a highly specialized technique, such as population surveys. Yet others help develop sample lists of stakeholders that OIOS-IED might otherwise have difficulty developing on its own. Whatever the specific purview of the evaluation advisory panel, it is good practice to develop an advisory panel ToR [IED #12] so that its members are clear on what is expected of them - and so that OIOS-IED profits from their expertise in an optimal manner. OIOS-IED mentions the use of advisory panels in the methodology section of its evaluation reports and provides a brief description of the panel, its function and membership.

3.   The OIOS-IED Quality Assurance System (QAS)

As Figure 3 illustrates, utilization of OIOS-IED's evaluations hinges on the credibility of the products it produces and the processes it follows. The OIOS-IED Quality Assurance System (QAS) was developed to ensure consistently high quality across all of its evaluations. Established in 2008 and revised in 2013, the OIOS-IED QAS includes ten elements. These are:

- Quality checklists that include minimum quality standards for each step of the inspection and evaluation process (checklists are referenced under each step of the evaluation process in Part II of this manual). Checklists are filled out by the Team Leader and certified by the Section Chief. The Directorate might also review them.

- Division-wide brainstorming sessions are held to ensure that evaluation teams benefit from the insights and experience of all OIOS-IED staff. Brainstorming sessions are mandatory for every project at the scoping stage and for the development of preliminary evaluation results. Teams may also request brainstorming sessions at other points in the evaluation process.

- Tours de table - i.e., brief updates from evaluation teams are scheduled at the end of Division meetings, with the opportunity to raise project challenges, issues and good practice with time-managed discussions among Division staff.

- Section Chief review and approval of project documents according to a schedule agreed between the Section Chief and evaluation team at the beginning of each new evaluation. Documents for review include inception papers, data collection instruments, summaries of data analyses and draft reports.

- Directorate review and approval of inception papers as well as draft and final evaluation reports.

- Section Chief and Directorate quality spot checks. For the Directorate, these are judicious, focusing on particular items underpinning critical results and recommendations. For the Section Chief, they include the sampling of work products at each step of the evaluation process – e.g., spot checks of interview notes, survey coding, document reviews, direct observation sheets, and analyses of data sources. Team members are expected to double-check their own work throughout and to assist in checking others' work.

- Fact checking of reports. This is conducted by an OIOS-IED staff member not on the evaluation team or an external source such as a consultant of all main numbers and results reported in draft evaluation reports to verify the accuracy of the information reported.

- Advisory panels (Part I, Section 2.3), established to provide input into evaluations at critical stages, such as during scoping, development of the inception paper, data collection and drafting of the evaluation report.

- External review of all final OIOS-IED evaluation reports. This is undertaken at the end of each biennium within the context of OIOS-IED programme performance reporting, to include both a review of the quality of its reports (technical quality and quality of engagement with evaluand/stakeholders) and a user feedback survey. In order to continually learn and improve from these peer review and client assessments, OIOS-IED has identified the ratings emanating from these two sources as indicators for measuring its progress against with the OIOS-IED PIP (Part I, Section 1.2).

- Periodic review and refinement of the OIOS-IED QAS.

The QAS clearly underlines the notion that quality is the responsibility of all OIOS-IED staff. In addition, OIOS's Self-Evaluation Policy [IED #5] commits the Office and its divisions to regular external evaluation. These independent exer-

cises shed additional light on the quality of OIOS-IED's work from an external perspective so that the Division can continually learn and improve.

Part II.   Steps of the OIOS-IED Inspection and Evaluation Process

Part II of the manual focuses on the practical aspects of OIOS-IED's work - that is, how it undertakes a typical inspection or evaluation (Part I, Section 1.4), step by step. Part II has two main chapters:

- Overview of the OIOS-IED Inspection and Evaluation Cycles (Part II, Chapter 1)
- Step-by-Step Guidance for Each Stage of OIOS-IED's Work Cycle (Part II, Chapter 2)

1.   Overview of the OIOS-IED Inspection and Evaluation Cycles

Although OIOS-IED follows a standard evaluation project management cycle, there are in fact two distinct tracks within this standard evaluation cycle - one for its RB-funded programme and thematic evaluations, and another for QSA-funded peacekeeping evaluations. Both types of evaluation span a 12-month period, but differ from each other in a number of ways.

Table 4 summarizes the differences between the RB and QSA work cycle.

Inspections are more ad hoc in nature, are not regularly planned, and are typically very narrowly focused. They therefore usually follow a shorter cycle than the 12 months allotted to OIOS-IED evaluations.

Table 4: Key Differences in the Work Cycle for Standard Evaluations and Peacekeeping Evaluations

		Regular Budget Evaluations	Peacekeeping (Support Account) Evaluations
Budget Source		Regular Budget	Support Account
Overall Time-frame	Project scoping and drafting and finalization of inception paper or ToR	April to June	March to May
	Data collection and analysis	July to November	June to October
	Report drafting	December to March the following year	November to February the following year
	Report finalization	End March	End February

OIOS-IED evaluation reports are typically submitted to the CPC or another inter-governmental body. Peacekeeping reports are submitted to the Fifth Committee of the GA. For both types of report, the Department for General

Assembly and Conference Management (DGACM), which processes all reports destined for the GA, assigns an official "slot date" around the timeframes indicated in Table 4. The slot date represents the official deadline by which time the report must be submitted to DGACM for editing and translation. Failure to meet the slot date results in an official designation of a late report. If reports are to be submitted to another inter-governmental body instead of, or in addition to, the CPC or Fifth Committee (e.g., Special Committee on Peacekeeping Operations, known as the C34), the cycle will likely be different.

For evaluation reports submitted to UN Secretariat programme managers rather than to an inter-governmental body, the final submission date similarly represents a deadline that is monitored by, and reported to the IAAC. Failure to meet the submission deadline also results in an official designation of a late report.

Evaluation cycles can be affected, and hence modified, by several different factors. Among these are the lack of a clear work plan, difficulties with evaluand cooperation, delays in mission planning, ad hoc requests for evaluations that have a different timeframe, and delays in the review process. In its PIP, OIOS-IED has identified these factors as issues to manage so that it delivers its evaluations in the timeliest manner possible.

2.   Step-by-Step Guidance for Each Stage of OIOS-IED's Work Cycle

While inspections and evaluations (Part I, Section 1.4) constitute different oversight tools, each shares the same basic steps. These are:

- Announcement (Part II, Step 1)
- Evaluation Design (Part II, Step 2)
- Data Collection (Part II, Step 3)
- Data Analysis (Part II, Step 4)
- Report Preparation (Part II, Step 5)
- Dissemination (Part II, Step 6)
- Post-evaluation "Housekeeping" Activities (Part II, Step 7)
- Tracking and Follow-up to Recommendations (Part II, Step 8)

These steps in the process are often iterative and overlapping: at any given point in time, two or more steps might be going on simultaneously. In addition, at some points there might be a return to a previous step. For example, one or more evaluation team members might need to undertake further data collection during the data analysis stage, or even the report preparation stage.

The remainder of Part II discusses each of these eight steps in detail and offers guidance on each. A dedicated sub-chapter is devoted to each step, with these sub-chapters indicated graphically using the sequential timeline bar provided below.

| Step 1: Announcement | Step 2: Evaluation Design | Step 3: Data Collection | Step 4: Data Analysis | Step 5: Report Preparation | Step 6: Dissemination | Step 7: Housekeeping | Step 8: Tracking Follow-Up |

Step 1:   Announcement

Step 1: Announcement → Step 2: Evaluation Design → Step 3: Data Collection → Step 4: Data Analysis → Step 5: Report Preparation → Step 6: Dissemination → Step 7: Housekeeping → Step 8: Tracking Follow-Up

The announcement stage marks the official launch of an OIOS-IED evaluation. This step is vital for conveying, at the very outset of the evaluation, the tone and "ground rules" of the evaluation to come. In addition to communicating to the evaluand the basic aspects of what OIOS-IED does, OIOS-IED also seeks to signal its balanced approach to the conduct of its work at this early stage - i.e., the balance of its mandated operational independence on the one hand, and its commitment to a consultative process on the other, so that its evaluations are used.

Once OIOS-IED has identified its work plan through its Strategic Risk Assessment Framework (Part I, Section 2.1), OIOS-IED's Director shares the full work plan by way of a memorandum [IED #13] to the USGs of all UN Secretariat programmes slated for evaluation during the forthcoming biennium.

From here, OIOS-IED launches each individual evaluation separately through a follow-up communication, the formal evaluation notification memo (Part II, Section 1.1), followed by an information brochure (Part II, Section 1.3) explaining OIOS-IED and its approach.

This sub-chapter has three main sections:

- Formal Announcement (Part II, Section 1.1)
- Preliminary Engagement with Evaluands (Part II, Section 1.2)
- Information Brochure (Part II, Section 1.3)

1.1   Formal Announcement

Subsequent to the OIOS-IED Director's memorandum sharing OIOS-IED's work plan for the forthcoming biennium, OIOS-IED's designated Section Chief and evaluation Team Leader informally contact the head of the evaluation unit of the UN Secretariat programme slated for evaluation to discuss preliminary issues of a technical and strategic nature. This provides an opportunity for the evaluand to inform OIOS-IED of critical factors potentially affecting the evaluation (e.g., major evaluations currently underway, recent change of leadership, and so forth). It also affords OIOS-IED the opportunity to request materials essential to its due diligence that are not publicly available.

Following this informal consultation with the evaluand - or, in the case of thematic evaluations, with the entity with overall responsibility for the policy (Part I, Section 1.4) - to discuss the overall timing and nature of the evaluation, and once the new evaluation has been formally assigned to an OIOS-IED evaluation team (Part I, Section 2.2), the Team Leader drafts a formal evaluation notification memorandum [IED #14], which OIOS-IED's Director sends to the USG or Head(s) of Department of the entity being evaluated. In addition to copying in the entity or entities concerned, OIOS-IED copies in the OIOS USG and the Director of the Internal Audit Division (IAD), as well as the BOA and the Joint Inspection Unit (JIU).

The notification memo includes a brief description of the evaluation, the evaluation team, and how OIOS-IED conducts its work. It also requests the programme's USG to nominate an evaluand focal point to work with OIOS-IED on the specific evaluation at hand. Attached to the notification memo is an aide mémoire [IED #15] that underscores OIOS-IED's mandate and other pertinent background information, and informs the evaluand of how the evaluation will progress.

1.2   Preliminary Engagement with Evaluands

Once the evaluand's USG, in response to OIOS-IED's formal evaluation notification memorandum (Part II, Section 1.1), has officially informed OIOS-IED of its selection of an evaluand focal point, OIOS-IED's designated Section Chief and the evaluation Team Leader arrange an official entry meeting with the focal point. This conversation allows the focal point to ask questions about OIOS-IED and its process and to provide OIOS-IED with important background information as it embarks on the evaluation. It also provides OIOS-IED an opportunity to explain its processes further, and to gain early insights relevant to the evaluation design stage (Part II, Step 2).

1.3   Information Brochure

OIOS-IED has developed an information brochure [IED #16] for informing a broad range of stakeholders about the work of the Division and specific evaluations. The brochure can be distributed at any time, but is generally used during the announcement stage described here and in the evaluation design stage (Part II, Step 2).

The brochure usually contains the following information:

- OIOS's and IED's background and mandate;
- The general purpose and objective of the evaluation;
- Its scope and proposed methodology (bearing in mind that it has yet to be scoped at this stage);
- Key evaluation questions (as above, to the extent any of this is known at this early stage);
- The evaluation timeline and type of report to be issued (GA or UN Secretariat programme managers); and
- Evaluation team members.

Step 2:   Evaluation Design

Step 1: Announcement	Step 2: Evaluation Design	Step 3: Data Collection	Step 4: Data Analysis	Step 5: Report Preparation	Step 6: Dissemination	Step 7: Housekeeping	Step 8: Tracking Follow-Up

The objective of the evaluation design stage is to define the contours of the evaluation - what aspects of the programme it will and will not look at, the questions it will ask within this scope, and how it will answer them - taking into account information of a strategic, practical and methodological nature that comes to light during this stage. The purpose of this stage, in keeping with OIOS-IED's targeted results as per its PIP (Part I, Section 1.2), is to set the evaluation up for success by ensuring that it is scoped to explore the most relevant issues, that the team producing it will be timely and efficient in delivering the evaluation, and that OIOS-IED and the evaluation itself will be credible both in product and process, so that the evaluation is ultimately used.

The evaluation design stage begins with a preliminary research stage (Part II, Section 2.1), coupled with on-going interaction with the evaluand and other stakeholders, in order to better understand the evaluand and its operating environment, and to determine the evaluation's scope, questions, methods and timeline. It culminates in the delivery of a high-quality inception paper (Part II, Section 2.9) (or in the case of inspections and many peacekeeping evaluations, traditional ToR) that speaks to these and other strategic and practical considerations as appropriate. In the interest of transparency, the inception paper also states how OIOS-IED arrived at its choices and why.

The inception paper serves as the central document anchoring OIOS-IED, the evaluand and other stakeholders in a shared understanding of how the evaluation will proceed.

The evaluation design stage differs for inspections and evaluations (Part II, Section 1.4) in OIOS-IED. The design of evaluations is lengthier and more complex than that of inspections and includes activities and outputs (such as the preparation of an inception paper) that do not apply in inspections. For this reason, this step focuses on design of evaluations only. In designing inspections, teams should use those aspects of this sub-chapter most relevant to the project at hand.

In explaining how OIOS-IED approaches the evaluation design stage, this sub-chapter has nine main sections:

- Undertaking Preliminary Research (Part II, Section 2.1)
- Conducting the Scoping Process (Part II, Section 2.2)
- Defining the Programme or Thematic Impact Pathway (Part II, Section 2.3)
- Selecting the Evaluation Topic (Part II, Section 2.4)
- Framing the Evaluation Questions (Part II, Section 2.5)
- Selecting Indicators (Part II, Section 2.6)
- Choosing the Most Appropriate Evaluation Design (Part II, Section 2.7)
- Planning Data Collection (Part II, Section 2.8)
- Writing High-quality Inception Papers (Part II, Section 2.9)

2.1   Undertaking Preliminary Research

After deciding to conduct an evaluation, putting together an evaluation team and announcing the project, OIOS-IED evaluation teams undertake preliminary research to familiarize themselves with the basics of the evaluand or theme at hand. Preliminary

research not only provides invaluable background reading on the evaluand. It also constitutes the due diligence that shapes all of the remaining aspects in the evaluation design stage described in this sub-chapter - and indeed, the data collection (Part II, Step 3) and data analysis (Part II, Step 4) stages to come.

Typical sources of information at this preliminary research stage include:

- S-G reports to the GA and/or UN Security Council [📖 S-G Reports];
- GA and UN Security Council resolutions [📖 Resolutions];
- Other UN governing bodies' official documents [📖 Official Docs];
- ST/SGBs on core UN Secretariat programme functions;
- UN Secretariat programme Strategic Frameworks;
- Budget proposals, fascicles and programme performance reports;
- Corporate policies and strategies;
- Integrated Monitoring and Documentation Information System (IMDIS) data;
- Other OIOS evaluation and audit reports on the programme/topic;
- JIU and BOA reports on the programme or theme; and
- UN Secretariat programme self-evaluation reports.

OIOS-IED usually requests these and any other relevant documents in the form of a requisition list [IED #30]. It tracks receipt of documents on an on-going basis. Throughout the evaluation process, this requisition list is often updated as further information and data needs emerge, and consolidated wherever possible in order to avoid overburdening the evaluand.

In addition to these internal UN sources, OIOS-IED may also consult external evaluations, reviews, studies and statistics at this stage. Such sources should be treated with care, however. Reputable think tank reports, journals, magazines or newspaper articles should receive greater weight than those less clearly reputable or less well known. Sources known to be of questionable integrity or bias are to be avoided altogether, unless there is a specific purpose for using them - e.g., if they

are useful to the evaluation team in shedding insight on key debates or controversies.

2.2   Conducting the Scoping Process

OIOS-IED's targeted result of timely, objective, credible and relevant evaluations that are used to improve UN Secretariat programmes is ambitious, particularly when viewed against the limited time (12 months) and resources (typically two to three staff) at its disposal. To ensure that it uses its time and resources efficiently, and that it adds real value through its evaluations, OIOS-IED invests in carefully defining the scope of the planned evaluation.

Responsibility for scoping lies with the Team Leader and the evaluation team. Section Chiefs provide supervision, feedback and guidance throughout the process. The OIOS-IED Quality Assurance System (QAS) (Part I, Chapter 3) includes a scoping checklist [IED #17].

The overarching goal of the scoping process is to delimit the boundaries of the evaluation - what it will and will not focus on - in light of key practical, methodological and strategic considerations. Examples of such considerations include the following:

- Strategic considerations - What seem to be the highest priorities and risks in the programme that an OIOS-IED evaluation could feasibly address? What major decisions by Member States and/or the evaluand stand to benefit from an OIOS-IED evaluation? What evaluative exercises have already been completed, are underway, or are planned that might affect the utility of the various scoping options being considered, and what is the capacity of the evaluand's self-evaluation function to satisfactorily complete them? What are the expectations of the CPC for the biennium - e.g., broadly or narrowly scoped evaluations?

- Methodological considerations - How high in the results chain can OIOS-IED measure effectiveness or impact (Box 3)? What is the overall evaluability of the various scoping options being considered - i.e., to what extent are key elements for a successful evaluation in place to feasibly pursue the various options (e.g., availability of data for answering foreseen evaluation questions, a clear shared understanding by the evaluand and its key stakeholders of the results being targeted by the programme or sub-programme in question, adequacy of resources to achieve the targeted results, realistic time to payoff of the programme in question compared to where it is in its implementation)?

- Practical considerations - What are OIOS-IED's time and resources, compared to the size of the evaluand, its programmatic breadth, and/or geographic coverage, and against the various scoping options emerging during this stage?

Additional goals of the scoping process include the following:

- Continued cultivation of rapport, positive engagement and credibility with the evaluand;

- Identification of potential strategies and opportunities for ensuring high utilization;

- Identification of high priority evaluation topics that should be addressed by either self-evaluation units or others (e.g., JIU);

- Determining whether OIOS-IED has evaluated the evaluand in the past eight years, and if so how the evaluation team will go about measuring the impact of its previous evaluation(s) of the evaluand, as per the commitment to measure its progress against the OIOS-IED PIP (Part I, Section 1.2);

- Articulation of a risk management strategy for addressing any factors that might hinder the timely completion of a relevant, objective and credible evaluation with the scope being proposed; and

- Other aspects that the Team Leader, in consultation with the Section Chief and team members, deems relevant.

INFORMAL PRELIMINARY CONSULTATIONS

Scoping actually begins with the in-depth background reading during the preliminary research stage (Part II, Section 2.1), coupled with preliminary discussions with the evaluand's head of evaluation and the evaluand focal point during the announcement stage (Part II, Step 1). A large amount of relevant information usually emerges during these stages that helps shape the evaluation scope.

It is at these two preliminary stages that the evaluation team begins contracting, and informally consulting with, subject matter experts and/or an advisory panel (Part I, Section 2.3), if one is engaged, to help inform its scope. It is also the time when the evaluation team reaches out to OIOS colleagues to help inform its thought process. This internal OIOS consultation entails coordination with the Internal Audit Division (IAD) and (less frequently) the Investigations Division (ID) to share information and avoid undue overlap. It also entails internal Division-wide brainstorming sessions, which the team undertakes early in the scoping process. These Division-wide brainstorming sessions, though informal, are a mandatory part of the scoping process. They are not intended to be formal presentations of the team's fully developed scope for colleagues' validation. Rather, they are meant to be a forum to help ensure that the team, in identifying its scope and overall approach, take advantage of the wealth of knowledge that OIOS-IED's diverse staff collectively embody.

SCOPING MISSIONS

From here, the scoping process becomes more formalized. In reality, it also begins to encompass the other areas covered in subsequent sections of this sub-chapter,

since these are indispensable to defining the scope. Evaluation teams might therefore decide to undertake short and targeted scoping missions to achieve the objectives of the scoping process and these other stages. In some cases, scoping missions can be undertaken virtually by video teleconference (VTC), phone, Skype and email.

The purpose of a scoping mission is different from that of a data collection mission, and evaluation teams need to carefully explain the difference to stakeholders, both in the information brochure (Part II, Section 1.3) and during the mission. Put simply, the questions OIOS-IED asks during the scoping mission are limited to those that help shape the contours of the evaluation. During the data collection mission, OIOS-IED seeks answers to the evaluation questions related to relevance, effectiveness, impact and efficiency.

BALANCING INDEPENDENCE AND CONSULTATION DURING SCOPING
Another communications challenge that can occur in engaging with stakeholders during this stage has to do with balancing OIOS-IED's mandated operational independence with its commitment to consultation. While OIOS-IED is open to any and all topics, including those that stakeholders are most interested in from a demand-driven standpoint, as an independent oversight body it must ultimately scope the evaluation to explore what most needs to be evaluated. It is important that OIOS-IED convey this nuanced point collegially but clearly during the scoping process. Through systematic analysis of relevant documentation and other information during the preliminary research stage, OIOS-IED independently identifies topics that address the highest priorities and risks that it sees for the evaluand, which it then discusses with stakeholders during the scoping process. This approach does not preclude OIOS-IED from maintaining an open dialogue with the evaluand on potential alternatives.

Whatever its selected topic, OIOS-IED takes a transparent approach to conveying the rationale for its selection in the inception paper (Part II, Section 1.1). Any topics considered but not selected for evaluation that OIOS-IED still considers to be of a high priority are also highlighted in the inception paper.

2.3   Defining the Programme or Thematic Impact Pathway
It is standard OIOS-IED practice to construct an evaluation Programme Impact Pathway (PIP) or Thematic Impact Pathway (TIP) [IED #18]. Although developed in parallel to the scoping process (Part II, Section 2.2), PIPs and TIPs warrant specific attention because of their importance to the evaluation.

For OIOS-IED's purposes, PIPs and TIPs are identical, differing only in the unit of analysis they reflect - i.e., a PIP portrays a single programme in its entirety, while a TIP portrays a specific subprogramme, policy, theme, or a joint initiative involving multiple programmes. OIOS-IED therefore uses PIPs mostly for broadly-scoped programme evaluations and TIPs for narrowly-scoped programme evaluations as well as thematic evaluations.

As with OIOS-IED's own PIP (Part I, Section 1.2), the PIPs and TIPs it develops for its evaluations represent visual roadmaps that convey what the programme (or subprogramme, policy, theme, and so on) is ultimately trying to achieve and how it envisions doing so. Their objective is to break the programme, policy or theme into its parts - and to reassemble the parts in order to convey the logic underlying the programme, policy or theme being evaluated.

PIPs and TIPs serve the purpose of helping set the evaluation up for success. They do so by ensuring OIOS-IED's understanding of the entity and thereby strengthening its credibility with the evaluand. They also contribute to the evaluation's success by anchoring the evaluation in a clearly shared understanding between OIOS-IED and the evaluand of what is meant by the relevance, effectiveness, impact and efficiency of the entity being evaluated. This is in part why OIOS-IED evaluation teams develop the PIP or TIP in consultation with UN Secretariat programme staff and other key stakeholders, in addition to conducting their own research.

In developing a PIP or TIP, the evaluation team answers a simple set of questions about the programme, policy or theme at hand. These questions are indicated in Figure 7.

The "control" triangle pointing to the right underlines the limited control that programme management often has as one moves along the various levels of the PIP.

This is an important consideration in addressing issues of internal and external factors in the evaluation, as well as issues of attribution or contribution.

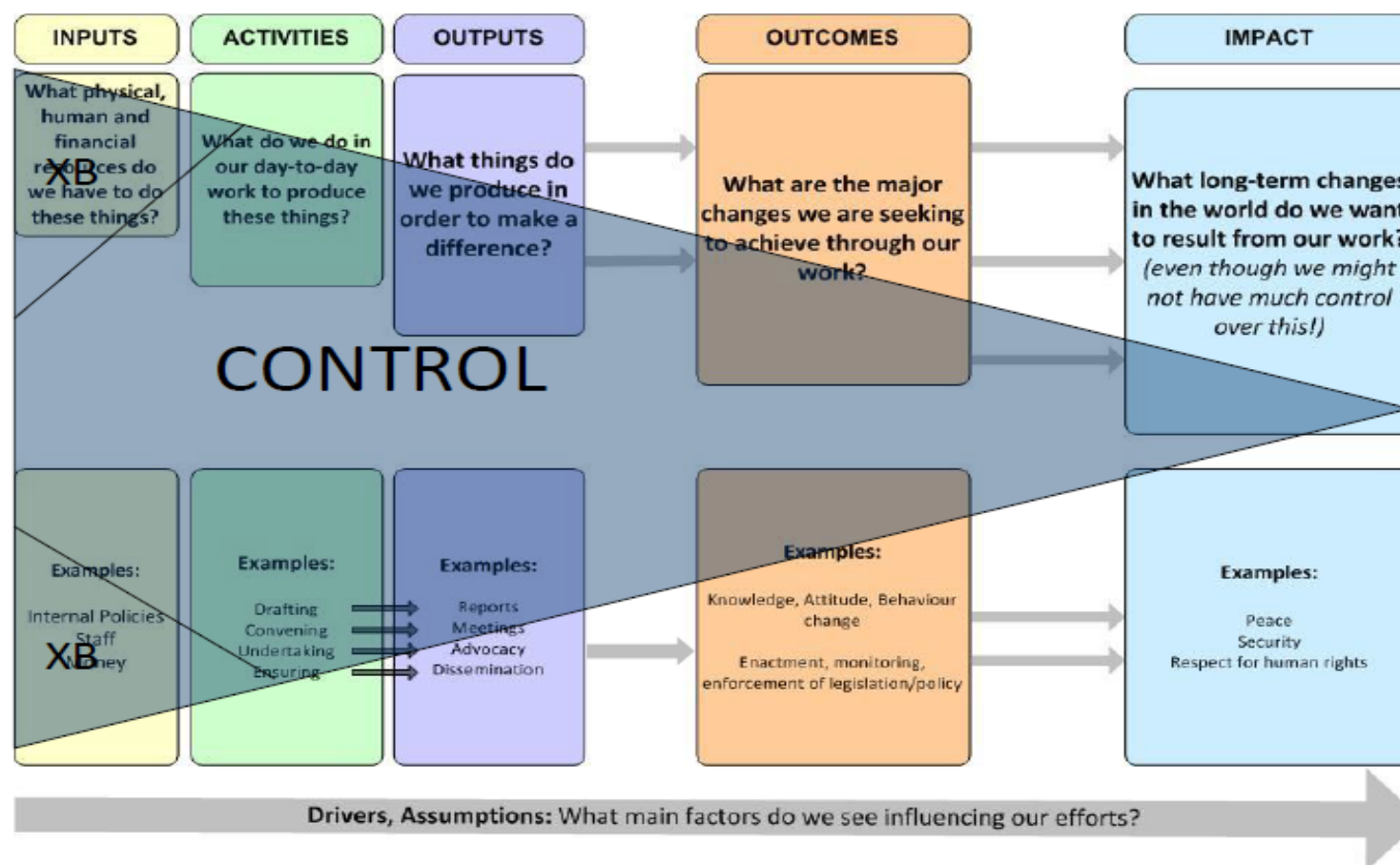Figure 7: Questions Asked at Each PIP/TIP Level, and the Programmes' Sphere of Control over Each Level

Figure 8 breaks down the questions OIOS-IED asks in developing the PIP or TIP
even more simply.

Figure 8: Questions Asked in Developing a PIP or TIP, Further Simplified
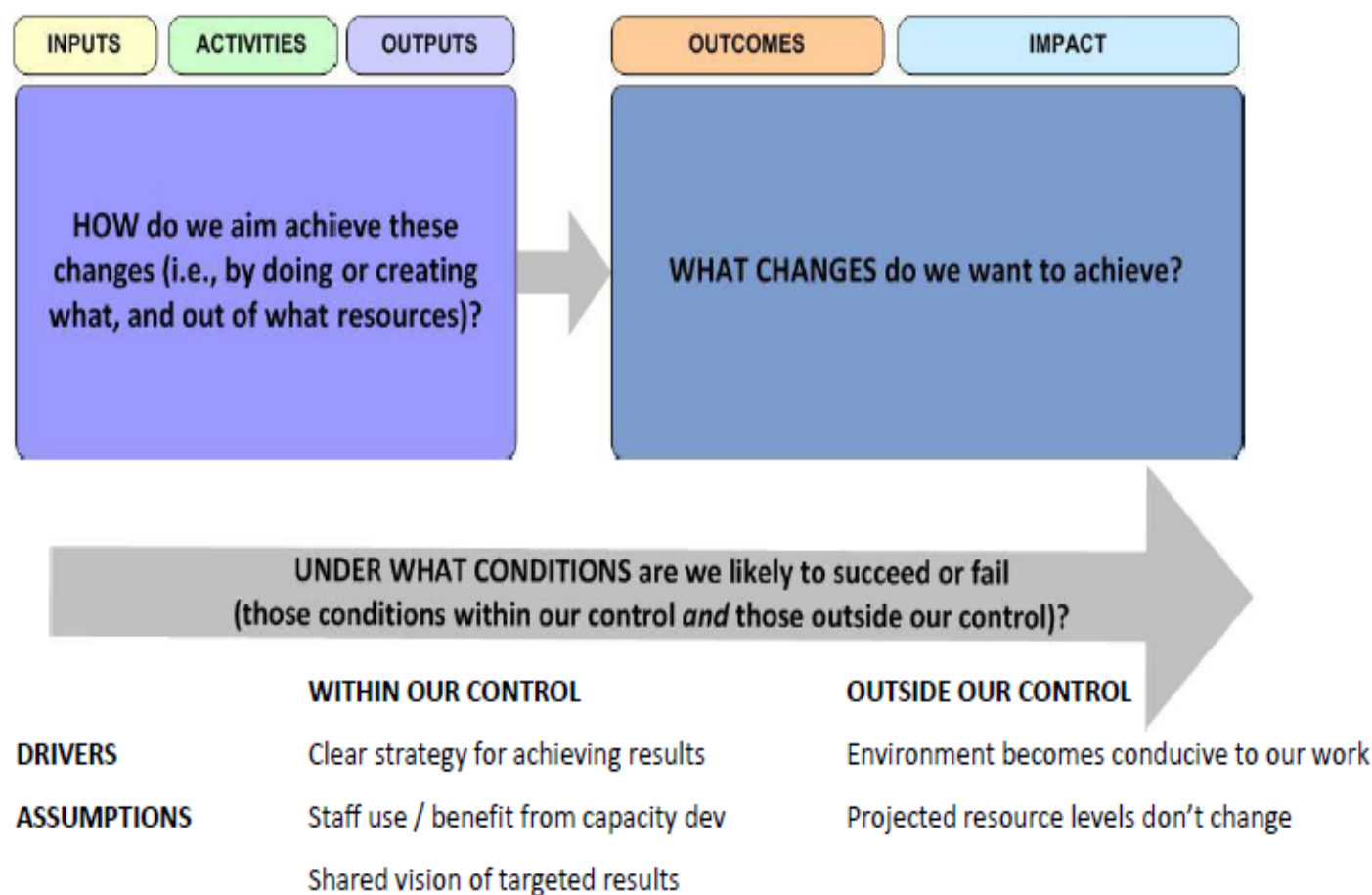
Figure 7 and Figure 8 show pathways linking the resources available to the programme (inputs) to what the programme does (activities) and delivers/produces (outputs) with these resources, and most importantly, toward what ends - i.e., the changes in conditions the programme seeks to achieve (outcomes and impact). These input-activity-output-outcome-impact linkages are often referred to as "results pathways."

Outcomes are often interlinked, either in parallel or sequentially in a series of dependencies, and are therefore often portrayed as stepping stones. Oftentimes these stepping stones are about influencing people, institutions, governments and others to take actions that will contribute to impacts such as a reduction in organized crime (UNODC), reduction in the impacts of disasters and complex emergencies (OCHA), reduction in gender-based violence (UN Women) or minimization of threats to the environment and their consequences to human wellbeing (UNEP). In any PIP or TIP, there might be several outcome pathways, each with several stepping stones on the way to the impacts. Arrows show the expected pathways amongst the various stepping stones in the diagram.

The PIP and TIP also include a number of statements about the conditions whose presence propels the UN Secretariat programme forward in achieving its goals (drivers) and the conditions without which such achievement is unlikely (assumptions). Some of these lie within the programme's control (internal factors) and others lie outside its control (external factors).

Despite the simplicity of the questions posed in Figure 8, the process of answering these questions in order to develop a specific PIP or TIP is not always straightforward. Oftentimes, even after extensive preliminary research (Part II, Section 2.1) and scoping (Part II, Section 2.2), the boxes contained in each level of the PIP or TIP, and especially the connections among them by way of the causal arrows, are not always clear or explicit.

Box 3: Where in the Results Pathway Does OIOS-IED Apply the Effectiveness or Impact Lens?

Depending on what is revealed during the preliminary research (Part II, Section 2.1) and scoping (Part II, Section 2.2) stages, OIOS-IED, in any given evaluation, will not necessarily address all of the evaluation criteria that it is mandated to assess. How it assesses effectiveness and impact can be particularly problematic.

Generally speaking, OIOS-IED strives to focus its assessment on the highest level in the results chain possible. When the timing, resources and evidence permit, OIOS-IED might design and carry out an impact evaluation [➔ UNEG. 2013b]. That said, the impact level is most often not evaluable, so the outcome level usually represents the highest-level achievement that can feasibly be measured. This is owed to the complexity of the programmes it evaluates, the inadequate state of valid and reliable data on the impact level, and OIOS-IED's own time and resource constraints compared to the intensive effort impact-level evaluations entail.

Identifying which level in the results chain the evaluation will focus on is one major benefit of developing a PIP or TIP during the scoping process. OIOS-IED uses the following logic in making this decision:
i) impact should not be looked for unless there is evidence of outcomes;
ii) outcomes should not be looked for unless there is evidence of outputs; and
iii) outputs should not be looked for unless there is evidence of activities being implemented.
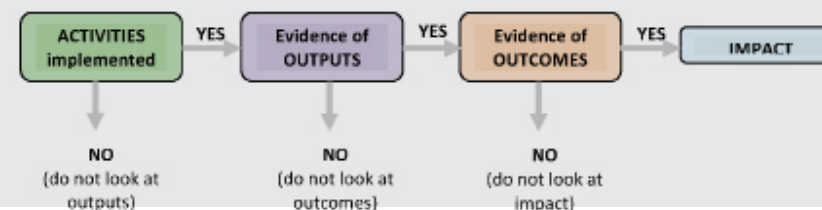
2.4 Selecting the Evaluation Topic

Having undertaken preliminary research (Part II, Section 2.1), collected scoping data (Part II, Section 2.2) and defined the PIP or TIP (Part II, Section 2.3), the next step is to identify all potential priority topics, and then to select one of these as the focus for the evaluation.

In identifying topics, both "vertically" and "horizontally" oriented alternatives are considered. Vertical topics are those that follow the organizational structure of the UN Secretariat programme. These might be the programme's organizational units - subprogrammatic and other divisions, units and sections, offices away from headquarters, field operation units, and so forth. Horizontal (or thematic) topics are those that cut across or are shared by several or all of the programme's organizational units. These might be activities, processes or results - e.g., programme management, policy and/or guidance, normative, analytical and/or operational work, or overall results in primary substantive areas (poverty reduction, sustainable development, gender, and so forth). Analysis of these topics should include a risk assessment, taking into consideration their relative size (in terms of staff and resources), the size and importance of their operational impact (e.g., in terms of clients or beneficiaries affected) and the risks to performance or results achievement in their specific contexts (in terms of volatility and complexity of the topic). While this list of potential vertical and horizontal topics cannot be exhaustive, it should be comprehensive, covering all activities and units of the programme.

With the list of potential topics available, the next step is prioritization and selection of the evaluation topic, which is by way of weighing the relative importance of the potential topics (ranked by the results of the risk assessment), and the scope of work required against the time and resource constraints of the evaluation team. Remaining topics on the list that are beyond the capability of OIOS-IED, or strategically better suited for another time or other evaluation body to conduct, should be listed in the inception paper (Part II, Section 1.1) and then referred to the concerned programme management, or to other relevant oversight or evaluation bodies as appropriate. They should also be reiterated in the scoping section of the final evaluation report, as as appropriate in the report's recommendations.

2.5   Framing the Evaluation Questions

Based on the evaluation topic and scope identified, OIOS-IED develops a series of evaluation questions to explore the major issues associated with this scope. Developing and refining evaluation questions (and sub-questions) are central to OIOS-IED evaluations, constituting the prerequisite for developing the evaluation indicators and methodology. Together, the evaluation questions, indicators and methodology form the basis of the evaluation design matrix [IED #19], which brings these elements together into a cohesive technical tool that guides the data collection (Part II, Step 3) and data analysis (Part II, Step 4).

OIOS-IED asks a combination of three different types of questions:

- Descriptive questions that determine "what is";
- Normative questions that compare "what is" with "what should be"; and
- Cause-and-effect questions that seek to determine "what difference" an intervention has made.

In framing evaluation questions, OIOS-IED evaluation teams think in terms of the following broad questions:

- Is the evaluand doing the right thing? How do they know? And how do they measure it?
- Is the evaluand doing these things right? How do they know? And how do they measure it?
- Is the evaluand doing these things on the right scale to make a difference? How do they know? And how do they measure it?

Figure 9 illustrates the difference between the three framing questions.

Figure 9: Framing the Evaluation Questions

Evaluation questions, whether in programme evaluations or thematic evaluations and regardless of how these are scoped, should be:

- Clear and precise, but - except in very rare cases - open-ended in nature;
- Clearly organized around the evaluation criteria OIOS-IED is mandated to assess - i.e., relevance, effectiveness, impact and efficiency (as well as any further OECD-DAC evaluation criteria OIOS-IED will be using in the evaluation);
- Presented in a logical order within the evaluation criteria;
- Directly and clearly grounded in the evaluation PIP or TIP (Part II, Section 2.3);
- The most critical for addressing the issues targeted by the selected evaluation scope; and
- Limited to a manageable number while allowing the evaluation to fulfil its accountability and learning objectives.

Furthermore, OIOS-IED evaluation teams should:

- Define effectiveness at the outcome or impact level, in relation to results achieved or not achieved in the entity and for the stakeholders targeted by the evaluand's work;
- Include "why" questions to identify key contributing factors (internal and external) that influence the performance of the evaluand or the theme/policy under evaluation;
- Include effectiveness/impact questions related to unintended as well as intended consequences;
- Include questions that specifically address human rights and gender equality (Part I, Section 1.6); and
- Include a question to ascertain the impact of previous OIOS-IED evaluation(s) of the evaluand within the past eight years (Part II, Section 2.2).

Before looking at the specificities of an individual evaluation, OIOS-IED evaluation teams are guided by a menu of generic key evaluation questions. These questions, summarized in Table 5, focus on programmes but can be easily adapted to thematic issues as well.

Table 5: Menu of Generic Key Evaluation Questions

Evaluation Criteria	Key Evaluation Questions
Relevance: The value added of the programme(s)	• What is the validity of the assumed input-output-outcome results chain (design question)? • To what extent does the UN Secretariat programme(s) fulfil different stakeholders' (e.g. Member States', targeted beneficiaries') needs and requirements, as articulated in the PIP and carried out in practice? • What is the congruence between GA mandates and the thematic or programmatic objectives of the work of the evaluand? • Do mandated objectives, proposed outcomes and outputs make sense in the current context, given changes since their design? • What is the level of satisfaction of key stakeholders with the (thematic) objectives and activities of the evaluand?
Efficiency: The timeliness and cost of the work of the programme(s)	• What is the timeliness/frequency/periodicity/timespan of production of outputs? • What financial and human resources (inputs) are required to produce outputs? • How do inputs compare with outputs? • How do productivity ratios compare with international comparators? • Are there lower-cost alternative strategies for contribution to outcomes? • To what extent do governance and management structures and processes (including coordination) enable or hinder delivery of products and services?
Effectiveness and Impact: The immediate outcomes achieved and the contribution made by the programme(s) in terms of impact associated with its/their work (long-term outcomes)	• To what extent are immediate outcomes shown in the PIP actually occurring? • If they are occurring, who/what is contributing to them (programme features/external factors)? • If they are not occurring, why not (programme features/external factors)? • If immediate outcomes are occurring, are they leading onto other outcomes and impacts in the PIP? • What unintended (positive/negative) outcomes might be occurring (because of the programme(s) and/or external factors)? • What is the magnitude of positive and negative outcomes that have actually occurred? • What is the level of satisfaction of different groups of key stakeholders? • What is the efficacy of partnership arrangements?

There are essentially four different types of work (Part I, Section 1.3) conducted by the UN Secretariat programmes inspected or evaluated by OIOS-IED - i.e., normative [➜ UNEG. 2013c], analytical, operational and internal support services. OIOS-IED adapts its evaluations to the different types of programmes. The key evaluation questions will reflect the nature of that work - i.e., normative (Table 6), analytical (Table 7), operational (Table 8) and internal support services (Table 9).

Table 6: Normative Work - Menu of Key Evaluation Questions

Evaluation Criteria	Key Evaluation Questions
Relevance	• In what areas is/are the programme(s) conducting normative work (as articulated in the PIP and carried out in practice)? • What roles is/are the programme(s) playing? • To what extent do(es) the programme(s) fulfil the needs and requirements of different stakeholders? • How sensitive is/are the programme(s) to gender equality and human rights in its/their normative work?
Efficiency	• What is the timeliness/frequency/periodicity/timespan of production of normative outputs? • What financial and human resources (inputs) are required to produce outputs? • How do inputs compare with outputs? • How do productivity ratios compare with international comparators? • Are there lower-cost alternative strategies for contribution to outcomes? • To what extent do governance and management structures and processes (including coordination) enable or hinder delivery of products and services?
Effectiveness and Impact	• To what extent are immediate outcomes shown in the PIP actually occurring? • If they are occurring, who/what is contributing to them (programme features/external factors)? • If they are not occurring, why not (programme features/external factors)? • If immediate outcomes are occurring, are they leading onto other outcomes and impacts in the PIP? • What unintended (positive/negative) outcomes might be occurring (because of the programme(s) and/or external factors)? • What is the magnitude of positive and negative outcomes that have actually occurred? • What is the level of satisfaction of different groups of key stakeholders? • What is the efficacy of partnership arrangements?

Table 7: Analytical Work – Menu of Key Evaluation Questions

Evaluation Criteria	Key Evaluation Questions
Relevance	• In what areas is/are the programme(s) conducting analytical work (as articulated in the PIP and carried out in practice)? • What role is/are the programme(s) playing in undertaking and disseminating analysis? • To what extent do(es) the programme(s) fulfil stakeholders' needs and requirements? • How sensitive is/are the programme(s) to gender equality and human rights in its/their analytical work?

Efficiency	• What is the frequency of production of analyses (outputs)? • How long does it take to produce analyses? Are outputs delivered in a timely manner? • What financial and human resources (inputs) are required to produce analyses? • How do inputs compare with outputs? • What is the unit cost of outputs relative to inputs? • Are there lower-cost alternative strategies for contribution to outcomes? • To what extent do governance and management structures and processes (including coordination) enable or hinder delivery of products and services?
Effectiveness and Impact	• To what extent are immediate outcomes shown in the PIP actually occurring? • If they are occurring, who/what is contributing to them (programme features/external factors)? • If they are not occurring, why not (programme features/external factors)? • If immediate outcomes are occurring, are they leading onto other outcomes and impacts in the PIP? • What unintended (positive/negative) outcomes might be occurring (because of the programme(s) and/or external factors)? • What is the magnitude of positive and negative outcomes that have actually occurred? • What is the level of satisfaction of key stakeholders? • What is the efficacy of partnership arrangements?

Table 8: Operational Work – Menu of Key Evaluation Questions

Evaluation Criteria	Key Evaluation Questions
Relevance	• In what areas is/are the programme(s) implementing operational activities (as articulated in the PIP and carried out in practice)? • What role is the programme(s) playing? • What is the congruence between GA mandates and the objectives of the programme(s)? • Are objectives in line with national priorities and international commitments? • Are they aligned with UNDAFs and other UN system joint strategies/programmes? • How sensitive is/are the programme(s) to gender equality and human rights in its/their operational work?
Efficiency	• What financial and human resources (inputs) are required to produce outputs? • Are outputs delivered in a timely manner? • How do inputs and outputs compare? • To what extent do governance and management structures and processes (including coordination) enable or hinder delivery of products and services?

Effectiveness and Impact	• To what extent are immediate outcomes shown in the PIP actually occurring? • If they are occurring, who/what is contributing to them (programme features/external factors)? • If they are not occurring, why not (programme features/external factors)? • If immediate outcomes are occurring, are they leading onto other outcomes and impacts in the PIP? • What unintended (positive/negative) outcomes might be occurring (because of the programme(s) and/or external factors)? • What is the magnitude of positive and negative outcomes that have actually occurred? • What is the level of satisfaction of key stakeholders? • What is the efficacy of partnership arrangements?

Table 9: Internal Support Services – Menu of Key Evaluation Questions

Evaluation Criteria	Key Evaluation Questions
Relevance	• What services do(es) the programme(s) provide to UN governing bodies (as articulated in the PIP and carried out in practice)? • To what extent do(es) the programme(s) fulfil stakeholders' needs and requirements?
Efficiency	• What financial and human resources (inputs) are required for providing services (outputs)? • Are services delivered in a timely manner? • How do inputs compare with outputs? • Are there lower-cost alternative strategies for contribution to outcomes? • To what extent do governance and management structures and processes (including coordination) enable or hinder delivery of products and services?
Effectiveness and Impact	• To what extent are immediate outcomes shown in the PIP actually occurring? • If they are occurring, who/what is contributing to them (programme features/external factors)? • If they are not occurring, why not (programme features/external factors)? • If immediate outcomes are occurring, are they leading onto other outcomes and impacts in the PIP? • What unintended (positive/negative) outcomes might be occurring (because of the programme(s) and/or external factors)? • What is the magnitude of positive and negative outcomes that have actually occurred? • What is the level of satisfaction of key stakeholders? • What is the efficacy of partnership arrangements?

2.6   Selecting Indicators

The second step in developing the evaluation design matrix (Part II, Section 2.5) is the selection of one or more indicators for each evaluation question. Indicators are either quantitative (numeric) or qualitative (narrative) variables meant to provide a clear, straightforward means of answering evaluation questions.

Mixing qualitative and quantitative methods, OIOS-IED strives to select a good mix of quantitative and qualitative indicators that are meaningful, unambiguous, disaggregated in order to explore disparities, manageable in number and practical to measure. Oftentimes, the evaluand can help select useful indicators.

Like the indicators themselves, the data collected and the analyses performed to measure these indicators is referred to as being either qualitative or quantitative in nature. Table 10 provides examples of each of these two indicator types, and the typical strengths and limitations of data associated with each.

Table 10: Strengths and Limitations of Quantitative and Qualitative Indicators

Type of Indicator	Strengths	Limitations
Quantitative indicators, e.g.: • Numbers • Percentages • Rates (e.g., mortality rate) • Ratio (e.g., sex ratio)	• Often sufficiently objective • Aggregation in order to establish magnitude of issues/sentiments is more feasible • Cross-country or inter-temporal comparison more feasible	• Picture that emerges is less rich and less nuanced than that obtained from qualitative data • Validity and reliability are often difficult (e.g. due to low response rates)
Qualitative indicators, e.g.: • "Compliance with…" • "Quality of…" • "Level of…" • "Degree of…" • "Satisfaction with…" • "The way that…" • "The nature of…"	• Measure judgments, opinions, perceptions and attitudes • Picture that emerges is richer (and more nuanced) than that obtained from quantitative data	• Usually highly subjective • Aggregation in order to determine magnitude of issues/sentiments is challenging • Cross-country or inter-temporal comparison challenging

One of the main rules of thumb for OIOS-IED in choosing indicators is to select the most suitable combination of quantitative and qualitative measures for answering the evaluation questions at hand. Rarely, if ever, is an OIOS-IED evaluation solely qualitative or solely quantitative in its analysis. OIOS-IED refers to this emphasis on using the best combination of qualitative and quantitative indicators as taking a mixed-method approach. A number of sources are available that provide strategies for strengthening an evaluation's mixed-method approach [➔ Rockefeller Foundation. 2012a] [➔ USAID. 2013a] [📖 MMIRA] [📖 BetterEvaluation] [📖 My M&E].

2.7   Choosing the Most Appropriate Evaluation Design

The third step in developing an OIOS-IED evaluation design matrix (Part II, Section 2.5), after choosing the evaluation questions and associated indicators, is to determine the overarching strategies for answering each evaluation question.

Evaluation literature refers to three broad categories of evaluation designs to collect and analyze data - experimental, quasi-experimental and non-experimental (descriptive) designs. There are many resources on these design options [➔ World Bank. 2009a] [📖 BetterEvaluation].

Because of the nature of the programmes and themes it evaluates, OIOS-IED most commonly applies non-experimental and quasi-experimental designs.

Experimental Designs

Experimental studies help provide evidence of a causal or correlational relationship between interventions and observed outcomes/impacts. They help the evaluator assess whether a desired result would have been achieved without a particular intervention. They are used to address cause-and-effect evaluation questions. Experimental studies involve random assignment of a representative number of individuals (or another unit of analysis) to either an experimental group (beneficiaries of the intervention) or to a control group (non-beneficiaries). While experimental designs are considered the optimum approach for excluding

the possibility that something other than a particular intervention led to an observed change, they are generally not feasible due to the type of interventions that OIOS-IED evaluates as well as time and money constraints.

Quasi-experimental Designs

Quasi-experimental studies can be used to obtain measurements before and after an intervention (such as the establishment of a peacekeeping mission) when it is not possible to randomly construct experimental and control groups. They are quicker and cheaper than experimental studies, but data reliability is lower. In quasi-experimental designs, individuals are assigned to so-called comparison groups based on a few essential characteristics (such as whether they lived in close proximity to a peacekeeping mission base). This involves the identification of a group of individuals assessed as being similar (comparable) to beneficiaries of an intervention, but who have not been exposed to the intervention. Changes in particular variables (such as exposure to armed conflict) might then be measured and compared in both groups.

OIOS-IED often makes use of time series designs, which, in their simplest form, take repeated observations of implementation of an intervention during the evaluation period, including at the end.

Non-Experimental (Descriptive) Designs

Non-experimental studies provide an in-depth description of a phenomenon or the relationships between two or more phenomena. They are generally used to answer descriptive evaluation questions. They help show whether or not UN Secretariat programme(s) and policies are operating as planned, provide feedback about services offered, determine whether or not the programme(s) and policies are producing the types of outputs and outcomes desired, and help clarify programme and policy processes, goals and objectives. They do not attempt to create intervention and non-intervention groups. Case studies (Part II, Section 3.1) are the main non-experimental design used in OIOS-IED.

2.8   Planning Data Collection

During the scoping process (Part II, Section 2.2), significant upfront thought must be given to the data collection stage of the project, which spans a large part of the evaluation cycle. OIOS-IED evaluation teams, in consultation with the evaluand, determine the type of data needed, identify where data are located (data sources) and agree on how best to retrieve them (data collection methods). This work is also summarized in the evaluation design matrix (Part II, Section 2.5). It is also important to identify in the evaluation design matrix the stakeholder groups that will be targeted in each data collection method.

To ensure the collection of high-quality and credible data, OIOS-IED evaluation teams collect a combination of different types of data originating from multiple data sources and using mixed data collection methods. OIOS-IED's mixed-method approach extends beyond blending a qualitative and quantitative approach. It also includes using the most appropriate data sources for answering the specific evaluation questions at hand. Some evaluation questions inherently entail qualitative methods (or specific data sources); others inherently entail quantitative methods (or specific data sources). Many, if not most, inherently entail both qualitative and quantitative methods, and a range of specific data sources.

Step 3 and Step 4 of OIOS-IED's work cycle delve more deeply into specific data collection methods and sources. It is useful at this stage, however, to briefly sketch out the overarching types of data sources OIOS-IED relies on to gather the evidence required to answer the evaluation questions at hand.

TYPES OF DATA SOURCES

Data can be classified in a number of ways. At the broadest level, data are either quantitative or qualitative in nature. In addition, data can be classified according to the specific source from which they are collected (e.g., survey data, interview data, and desk review data). Furthermore, data can be categorized into primary data and secondary data. Primary data are data collected by the evaluators themselves, while secondary data are data previously collected by others. Whether to use primary or secondary data is both a methodological and practical consideration. Using already existing secondary data, e.g., self-evaluations or IMDIS data, can save evaluators time and effort. Table 11 summarizes the advantages and disadvantages of each type.

Table 11: Advantages and Disadvantages of Primary and Secondary Data

Type of Data	Advantages	Disadvantages
Primary Data	• Within evaluators' direct control • More likely to meet the needs of the evaluation • Provide more recent information	• Time- and resource-intensive • Primary data collection oftentimes neither feasible nor cost-effective
Secondary Data	• Lower costs - No need to bear the expense of the collection, which can be a sizeable sum • Time savings - Obtaining and processing secondary data takes less time than collecting comparable primary data.	• Uncertain validity and reliability • Data might not reflect exactly what the evaluator would have chosen in terms of the population and variables when collecting new data • Data tend to be inflexible - It is not possible to revert with follow-up questions or to re-interview stakeholders

OIOS-IED data collection follows two basic rules of thumb. First, it uses secondary data whenever possible - i.e., wherever it is available and is sufficiently valid and reliable. Second, it collects only the amount and type of primary (first-hand) data needed to answer the evaluation questions.

Beyond this typology of primary and secondary data, the types of data needed are ultimately determined by the indicators selected. OIOS-IED generally relies on four types of evidence:

- Physical - obtained through direct observation of people, property or events. Such qualitative evidence is collected from primary sources in the form of written notes, photographs, drawings, charts, maps or physical samples;

- Documentary - qualitative or quantitative data in the form of written or other visual information, such as memos, reports, financial records, photographs, videos and other primary/secondary data source documents. Documentary evidence can be in electronic or hard copy format;

- Analytical - includes computations, comparisons and rational arguments; and

- Testimonial - obtained through interviews and surveys, testimonial evidence is particularly useful in identifying causal relationships.

STAKEHOLDER MAPPING
The scoping process should also include the development of a list of relevant stakeholder groups for the UN Secretariat programme or thematic issue being assessed. It is helpful to develop a stakeholder map [IED #20] to establish the relationships surrounding the evaluand. Stakeholder maps help facilitate the process of completing the "data sources" column of the evaluation design matrix (Part II, Section 2.5) as well as the process of choosing a sample of stakeholders from whom to collect information.

In addition, in keeping with OIOS-IED's focus on the utilization of its evaluations, the stakeholder mapping process helps evaluands themselves identify individuals to serve on any internal or external reference groups that the evaluand may choose to establish.

2.9   Writing High-Quality Inception Papers
Clear, crisp and concise writing is the hallmark of OIOS-IED's work, beginning during the evaluation design that culminates in the production of an inception paper for programme evaluations. OIOS-IED's Quality Assurance System (QAS) (Part I, Chapter 3) includes a checklist for inception papers [IED #21].

Inception papers are not prepared for inspections, nor are they always prepared for peacekeeping evaluations (Part I, Section 1.4). In these cases, traditional ToR [IED #22] will suffice.

Inception papers include several additional items not traditionally included in a ToR:

- An assessment of the evaluation capacity of the evaluand (or, in the case of thematic evaluations, the mandated or actual reporting arrangements for the particular policy or theme);

- The rationale for the selection of the evaluation topic; and

- Other topics warranting evaluation that OIOS-IED is unable to undertake.

There are any number of resources on results-driven writing that OIOS-IED relies on in drafting its inception papers and other products [➔ UNEG. 2010a] [➔ United Nations. 1984a] [➔ World Bank. 2011a] [📖 BetterEvaluation], including guides specifically geared to writing within the UN system.

Inception papers are typically anywhere from 8 to 15 pages long, depending on the

contours of the specific evaluation at hand. They are prepared by the evaluation team with feedback, guidance and review/approval from the OIOS-IED Section Chief. They should be clearly written, with a logical flow within and among sentences and paragraphs. Inception papers include the following sections:

- Background on the evaluand, including the PIP or TIP (Part II, Section 2.3) and M&E function and capacity;
- The process OIOS-IED followed in arriving at the choices it has made;
- The objective(s) of the evaluation;
- The scope of the evaluation (what it will and will not covered and why);
- Any previous evaluations, reviews or audits;
- Opportunities for future inspection or evaluation work;
- The evaluation issues (evaluation questions);
- The evaluation design matrix (Part II, Section 2.5);
- Any case studies selected, and the process and rationale behind this selection;
- Stakeholder mapping (Part II, Section 2.8);
- Mainstreaming of human rights and gender (Part I, Section 1.6);
- Use of consultants (Part I, Section 2.3);
- An evaluation risk management strategy, including working arrangements with the evaluand;
- A utilization strategy - i.e., potential decisions and actions the evaluation might influence, whether one or more reference groups are recommended, and so on;
- The evaluation timeline; and
- Resources required.

After final review by the Section Chief, inception papers are submitted to the OIOS-IED Directorate for review. They are subsequently shared with the evaluand and, if relevant, the advisory panel (Part II, Section 2.3), for comments. Evaluation teams must give fair consideration to the comments received and incorporate them as appropriate, but are not obliged to incorporate all suggestions made. The inception paper is then submitted again to the OIOS-IED Directorate for final approval.

Step 3: Data Collection

Step 1: Announcement	Step 2: Evaluation Design	Step 3: Data Collection	Step 4: Data Analysis	Step 5: Report Preparation	Step 6: Dissemination	Step 7: Housekeeping	Step 8: Tracking Follow-Up

Flowing directly from the evaluation design stage (Part II, Chapter 2) the objective of the data collection stage is to collect the best evidence available, using a mixed-method approach [📖 BetterEvaluation], for answering the evaluation questions. Its purpose, therefore, is to maximize the technical credibility of the ensuing data analysis (Part II, Step 4) and report drafting (Part II, Step 5), both of which hinge directly on the quality of the data collected during this stage.

This sub-chapter has three main sections:

* Data Collection Methods, Instruments and Processes (Part II, Section 3.1)
* Choosing Data Collection Methods (Part II, Section 3.2)
* Sampling (Part II, Section 3.3)

The OIOS-IED Quality Assurance System (QAS) (Part I, Chapter 3) includes a checklist for data collection [IED #23]. This checklist helps OIOS-IED evaluation teams ensure that critical items described in this section are included in the data collection plan.

3.1 Data Collection Methods, Instruments and Processes

This section provides an overall description of the data collection methods OIOS-IED uses, the relative advantages and disadvantages of each, and practical guidance on good practice when using each method.

The data collection methods OIOS-IED most frequently uses include:

* Interviews
* Focus groups
* Self-administered surveys
* Field-based surveys, also known as population surveys
* Direct observation
* Desk review, also known as document review
* Case studies
* Field missions

Less typically, OIOS-IED uses:

* Independent expert assessments
* Remote monitoring

The specific combination of these methods that OIOS-IED uses depends on several factors:

* The evaluation questions and their corresponding indicators and data sources, as shown in the evaluation design matrix (Part II, Section 2.5);
* The human and financial resources allocated to the evaluation;
* The time available for data collection; and
* The availability of data and the type of data available.

OIOS-IED will have already considered these factors during the evaluation design stage (Part II, Step 2) of the evaluation, and articulated its choices in the inception paper (Part II, Section 2.9). That said, it is not uncommon for some questions surrounding these factors to linger once data collection begins, particular the factor of data availability. In this way, OIOS-IED often has to reassess its choice of data collection methods well into the data collection stage.

Where relevant, the evaluation team develops sampling strategies, sampling frames, selection criteria and sampling techniques for each data collection modality and draws its samples (Part II, Section 1.1) accordingly. To ensure ready availability of data when data analysis (Part II, Step 4) and report drafting (Part II, Step 5) begins, OIOS-IED evaluation teams must agree on a system for keeping track of in-coming data from various sources (e.g., physical recording devices such as index cards or electronic tools such as simple spreadsheets).

INTERVIEWS
What they are and how OIOS-IED uses them
Interviews are conversations between the evaluator and stakeholder, involving an interviewer administering questions to one (or more) persons resulting in qualitative information. There are a number of external resources that summarize good practice for interviews [➔ Brinkmann and Steiner. 2014] [➔ World Bank. 2009a] [📖 BetterEvaluation].

The main advantages of interviews include:

- Gaining in-depth information and understanding of the evaluation subject and identifying key issues;
- Collecting different attitudes, opinions and perceptions from a wide range of stakeholders;
- Comparing and contrasting stakeholder perspectives;

- Collecting non-perceptual data to be verified later through other means; and
- Identifying additional data sources.

The main limitations of interviews include:

- The amount of time required for organizing and conducting them;
- The limited time the interview allows for capturing data on all of the evaluation questions at hand;
- The time and resources needed to compile and analyze data collected from interviews;
- The potential for interviewees to hold back in sharing their views candidly in a face-to-face, non-anonymous environment, particularly if the subject matter is sensitive or if they fear their confidentiality will be compromised;
- The potential for what is termed "contamination" among interviewees (i.e., when those interviewed earlier in the evaluation communicate the questions to those interviewed later, thus giving the latter group time to anticipate and prepare responses in a way that biases responses); and
- The risk of influencing responses through poor interviewing techniques.

OIOS-IED conducts three main types of interviews, as described in Table 12.

Table 12: Comparison of Interview Methods

Structured interviews	Structured interviews are based upon a standardized set of questions following a pre-determined order (formal). There is little or no space for improvisation. Structured interviews facilitate greater comparability of data collected.
Semi-structured interviews	Semi-structured interviews cover broadly the same set of questions as structured interviews, but not necessarily in the same pre-determined order. They provide a good balance between rigour and flexibility, as spontaneous questions are possible.
Unstructured interviews	Unstructured interviews have no predefined set of questions and are similar to a conversation on a particular topic. Questions are spontaneous.

OIOS-IED typically conducts these different types of interviews in three different modes - individual face-to-face interviews, which include executive interviews; individual or group telephone/Skype interviews; and face-to-face group interviews. Table 13 describes each of these.

Table 13: Three Different Interview Modes

Interview Modes	# of Interviewees	Description	Situation When Used	Main Advantages	Main Limitations
Individual face-to-face interviews	1	Perhaps the best known. The interviewer is in control of the event, directing the pace of the interview, as well as providing clarification to items in the interview guide and responses to these items, if necessary.	Used in almost all evaluations, individual face-to-face interviews are the standard means of obtaining valid and reliable perceptual data from individual stakeholders.	Adds to the overall validity of the data collected as there is less room by the interviewee to interpret the questions differently from their intent. Face-to-face interaction helps build rapport.	Often challenging to cover all material in a short period of time. Large number of such interviews requires careful and often time-consuming coding.
Executive interviews	1	A variation of the individual interview is the executive interview, an interview with a high-level official, e.g., Permanent Representative to the UN.	Usually conducted in order to ensure that this all-important voice is heard in the evaluation. Requires special planning and consideration of the individual's position and time, however. It is good practice to be prepared to restrict the interview to the most important issues/questions should the high-level official be pressed for time. OIOS-IED management should invariably be informed in advance of planned executive interviews to enable their attendance during the interview.	Gives voice to senior-level officials. Also enables evaluation team to attune them to the evaluation, if they have not yet been involved, and to obtain their views on the evaluation's potential uses and other strategic considerations.	Sometimes yield data of limited analyzability, owing to interviewee's lack of connection to day-to-day operations, and to interviewee's numerous political considerations.

Group interviews	2 or more	Similar to one-on-one interviews, the main difference being the number of interviewees whose perspectives are sought. Differs from focus groups in the number of interviewees (i.e., typically number fewer than 5), and in that dialogue and interaction among interviewees themselves is not a primary or explicit goal.	Usually undertaken out of convenience – i.e., when time or scheduling limitations do not allow separate interviews – or when a small number of staff members serve an identical function and/or are known in advance to have very similar perspectives.	Provides a more economical means of achieving wider coverage in a limited period of time. Though different from focus groups, group interviews do sometimes enable observation of group dynamics and insights into the interviewees´ behaviours and attitudes.	Places even greater pressure on interviewer to cover all material in a limited time. "Tour-de-table" or "round robin" response format can at times seem awkward.
Telephone/Skype interviews	1 or more	Not a separate interview type in its own, but rather a different (i.e., virtual) format for conducting any of the interview types above.	As it is not always practical to conduct face-to-face interviews with all stakeholders given the cost and environmental impact of travel, telephone or Skype interviews are standard procedure in most, if not all, OIOS-IED evaluations	Provides cost-effective alternative to in-person data collection, particularly when the limited number of stakeholders in the location does not warrant a full-scale mission.	Lack of face-to-face interaction, coupled with technical challenges, can limit free flow of exchange as well as the establishment of rapport.

Planning interviews

An interview plan ensures that information obtained will be of high quality. Interview plans should:

- Identify interviewees within each stakeholder group identified in the stakeholder mapping;
- Determine the mode(s) of interview to be conducted with each interviewee, as described in Table 13;
- Determine the interview method (structured, semi-structured or unstructured), as described in Table 12;
- List interview questions and instruments that will be used to record answers;
- Define the time frame for completing all interviews;
- Establish responsibilities and procedure(s) for contacting interviewees and conducting interviews;
- Assign OIOS-IED team members to the specific interviews where they can add the most value (e.g., by virtue of language ability, substantive experience, seniority), wherever this is important;
- Take into account individual professional development and management goals, as outlined in the team compact (Part I, Section 2.2);
- Describe post-interview review procedures; and
- Discuss arrangements for call backs to interviewees for verification or clarification purposes.

Determining whom to interview is critical to gathering a wide range of views, which in turn enhances the robustness of evaluation results. In addition to ensuring inclusion of key individuals within each group identified in the stakeholder mapping (Part II, Section 2.8), it is generally desirable to obtain a good representation of interviewees covering different gender, geographical representation and stakeholder perspectives. In the interest of ensuring optimal independence, wherever possible, OIOS-IED attempts to identify those individuals it would like to meet. That said, evaluand focal points and UN Secretariat programme staff whom OIOS-IED engages with during the scoping stage (Part

II, Section 2.2) of an evaluation can be very helpful in identifying potential interviewees.

OIOS-IED evaluation teams should thoroughly pre-test interview guides to determine whether the questions actually work, especially those regarding sensitive issues, and also to validate the order in which questions are asked (for structured interviews).

In terms of contacting interviewees and scheduling interviews, it is helpful to use an interview log to keep track of contacts and schedules. To enable comparison between information obtained by different interviewers, before beginning interviews, evaluation team members should meet to reach a common understanding of the interview questions and agree on a common introduction, neutral probes and responses to anticipated questions from interviewees.

Conducting interviews

The data collection tool for conducting interviews is called an interview guide. OIOS-IED has sample interview guides in place to draw on [IED #32 - 33]. Generally speaking, when developing interview guides, OIOS-IED staff members apply the following guidelines:

- Preface the guide with the main preliminary points that must be covered - a sincere thank-you, a brief background on OIOS, IED and the evaluation, a statement of how long the interview will likely take, a clear communication of OIOS-IED's confidentiality guarantee, and so on;

- Before moving onto the main interview questions, ask the interviewee an "ice-breaker" question - e.g., how long s/he has been working for the evaluand, what his or her role is;

- Make sure the questions are aligned to the evaluation questions;

- Ensure that, as with evaluation questions, the interview questions are open-ended in nature, in order to capture the largely perceptual, qualitative data typically being sought;

- Avoid loaded and leading questions;

- Keep the question language as clear and simple as possible, with the least amount of jargon possible. If technical terms are unavoidable, be sure to have a clear definition, to be used identically for all interviewees, should an interviewee be unfamiliar with the term;

- Sequence the questions in a way that promotes a natural conversational flow with clear, logical transitions;

- Clearly demarcate any intended pre-identified probes in the guide, and make sure instructions are clear for interviewers with regard to how to use the probes; and

- For any questions targeted to a specific stakeholder group, consider developing a separate interview guide document that includes these (as well as all applicable core interview questions), for ease of navigation for the interviewer.

There are several good practices to follow when conducting interviews. A shortlist of good practices is as follows:

- When beginning the interview, introduce OIOS-IED and explain the evaluation objective;

- Always assure interviewees of the confidentiality of the interview. This is another important element to soliciting optimal candour;

- Take time to establish and maintain positive rapport with the interviewee. Where culturally and socially appropriate, make eye contact and smile. Throughout the interview, acknowledge his or her responses with nods and other forms of neutral body language;

- Be an active listener, tuning out other distractions so that the interviewee feels s/he has your full attention and is being heard. Process the information the interviewee is offering, and occasionally mirror, synthesize, and where appropriate juxtapose his or her various points, prefacing your reply with phrases such as, "Just so I capture your views on this correctly, perhaps I can summarize?" or "I don't want to put words in your mouth, so may I quickly recap what I've heard you say?" Another option for clarifying interviewee responses is to ask the same question again;

- Keep track of time in order to cover all questions within a reasonable time, but do so discreetly, rather than making the interviewee feel rushed;

- Avoid leading or loaded questions, and avoid re-framing the open-ended questions formulated in the interview guide as yes-no questions;

- Always be neutral and polite. Use the neutral probes included in the interview guide, but only after allowing the interviewee to respond on an unaided or "top-of-mind" basis (i.e., without prompting). Keep an open mind and never appear to approve or disapprove of a response. If a response is ambiguous, find a neutral way of probing for a more definitive answer. One good way for addressing ambiguities is to ask for an example, or to ask "Can you discuss that further?" or "Can you give me an example of that?";

- Never express your own opinion, and never lead the interviewee to your own opinion as an interviewer - e.g., by interrupting the interviewee, completing his or her sentences with your own thoughts, or by mirroring the interviewee's response with transitional phrases such as "I suppose you mean...";

- Never debate or argue with the interviewee;

- If the conversation veers off topic, re-direct the interviewee back to the question in the guide, but do so politely. Exercise patience if interviewees are not focused or in some other way slow down the interview;

- Handle difficult interviews tactfully. Some interviewees might appear shy, bored or even hostile. Proceed with the questions as tactfully as possible and avoid further alienating the subject. Alternatively, owing to the sensitive nature of some topics or relationships, on occasion it is possible that an interviewee might become emotional. In these cases treat the interviewee with respect, compassion and empathy. No data are worth causing undue distress to an interviewee. Give him or her a moment and, as appropriate, offer to move onto another question or reschedule the interview;

- Be sensitive to the burden placed on the interviewee more generally. If time becomes a constraint, be strategic and concentrate on the most important questions. Flexibility and focus are essential;

- If more than one person is being interviewed at the same time (group interview), make every effort to allow each person to respond to every question being asked; and

- If a structured interview is selected, do not change the general sequencing of questions. Once the interview begins, it is important that all interviewees receive the same questions and in the same order. Sometimes, however, it is not possible to ask questions in a set order for various reasons, including shortage of time. In such circumstances, determine beforehand the most important questions that need to be addressed.

Box 4: Tips for Conducting Telephone or Skype Interviews

Telephone and, increasingly, Skype interviews are standard practice in most, if not all, OIOS-IED evaluations. Some additional specific elements that can enhance the effectiveness of telephone interviews are as follows:
- Set a date and time. Plan an exact time to have the interview and follow through with it;
- Factor in time zone differences. Online tools [📖 timeanddate] are available to help calculate accurate local times;
- Determine who will call whom. Typically, the OIOS-IED evaluator calls the interviewee. In some cases, however, interviewees prefer to initiate the call;
- Inform interviewees if they are on the speakerphone as well as about any other person in the room;
- If using Skype, test audio and video settings beforehand so as to avoid technical glitches;
- Ensure privacy;
- Be aware of the unique challenges involved in conducting interviews through these modalities. These include the following:
  - Audibility - audibility is often a particular challenge, owing to echoes, static, time lags in between questions and responses that result in unintentional interruptions, and more. If audibility seriously impedes the quality of the interview, suggest re-initiating the call and/or trying another modality. For Skype interviews, turning off the video feature often resolves these audibility problems;
  - Rapport - telephone and audio-only Skype interviews remove the face-to-face inter-personal element from the interview, making it more challenging to establish the rapport so conducive to candour. In these cases, establishing rapport using verbal rather than physical cues becomes particularly vital – e.g., by demonstrating extra patience with audibility problems, by clearly stating (and, where necessary, patiently re-stating) the questions, being especially vigilant about using a friendly tone, and so on; and
  - Agility - telephone and Skype interviews pose particular challenges to interview note-taking. It is advisable that headsets be used in order to free up both hands for note-taking if there is only one team member conducting the interview.

OIOS-IED does not typically audio-record interviews. Therefore, good note taking is needed to ensure the completeness and quality of data obtained during the interview process. In OIOS-IED, the preferred method is to type interview notes. It is important to transcribe as much of the interview as possible. The notes should never be taken in the third person (i.e, "he said that …") but rather should be recorded verbatim. Direct quotes should be identified in quotation notes. OIOS-IED is also exploring a number of technologies to enable tablet-based note-taking [📖 iPad] [📖 WritePad], which can be automatically transcribed, thus helping OIOS-IED achieve both efficiency and good interviewee rapport.

Post-interview follow-up
Once interviews are complete, and prior to their analysis (Part II, Step 4), some post-interview follow-up is required. The evaluation team should:

- Review each interview and make sure that all questions are answered. If not, it might be necessary to re-contact individual interviewees to complete the information;

- If notes have been handwritten, transcribe these;

- Clean the interview notes as soon as possible following the interview and store typed interview notes on the OIOS-IED shared drive in the relevant project folder;

- Consider how to handle non-responses (Part II, Section 1.1). Options include conducting a non-respondent analysis to determine how different non-interviewees are from interviewees, and thus the extent to which the existing answer set is gravely affected by the missing set of interviews. A more aggressive approach is to establish a statistical basis for imputing responses for each missing person ("imputation"); and

- Store raw interview notes and summaries in a secure location, accessible only to the evaluation team, in order to protect interviewees' confidentiality;

- Although not mandatory, consider sending a thank-you email to interviewees.

FOCUS GROUPS

What they are and how OIOS-IED uses them
OIOS-IED uses focus groups in its data collection to gain a more nuanced and richer understanding of the evaluation issues than individual interviews alone might allow. Although focus groups are similar to interviews in the questions they ask, their objective is different from that of interviews and their dynamics considerably different.

There are a number of external resources that summarize good practice for focus groups [➔ World Bank. 2009a] [📖 BetterEvaluation].

A focus group is a facilitated conversation, framed around the evaluation questions, which aims to gain greater insight into how people think about a specific issue or topic and why, and to understand behaviours and motivations. Its main objective is to arrive at a deeper understanding of the programme or theme at hand by encouraging participants to talk to each other rather than to the evaluator moderating the discussion. In this sense, focus groups differ from group interviews.

The main strengths of focus groups are that they:

- Give greater insight into how people think about a specific issue or topic and why, and assist in understanding behaviours and motivations;
- Solicit a range of opinions and perceptions (and not always consensus);
- Help separate fact from opinion (in that, through talking with each other, group participants often arrive at a consensus of what is the factual truth and what is not, or at least what is a matter of subjective interpretation); and
- Initiate a creative process that can help generate ideas for recommendations.

The main limitations of focus groups are that they might:

- Require more time for organization and conduct;
- Be more challenging to facilitate than one-on-one interviews;
- Cannot be generalized to the broader population of the focus group participants;
- Distort individual opinions due to group dynamics; and
- Result in open conflict that can be difficult to manage, and that has negative reverberations for the programme after the evaluators are gone.

Planning focus groups
The logistics of planning focus groups is virtually identical to that for interviews. Almost all of the guidance provided on interview planning therefore applies to focus groups as well.

Beyond these logistical issues, however, additional forethought is necessary in planning for focus groups, owing to the group dynamics that shape them and that they are intended to uncover. First and foremost, planning for focus groups requires careful consideration of exactly what data the evaluation team wants them to generate. Several points of guidance can help set this mode of data collection up for success. This includes the following:

- Selection of focus group participants should be done systematically, with a clear sense of how many focus group sessions are desired and according to a set of specific criteria. Composition should be governed by the purpose of the focus group and the type of information to be obtained. Participants should be sufficiently homogeneous to ensure that participants are able to engage in a common discussion, but with sufficient variation in views and opinions as much as this can be foreseen;

- The recommended group size is around five to eight persons. If fewer than this number, the focus group will more likely resemble the dynamics of a group interview. If more than this number, the focus group will likely become unwieldy to manage;

- Lasting roughly 90 to 120 minutes, focus groups are typically longer than interviews in order to allow all topics to be covered and all opinions to be heard;

- As discussed previously, different data collection methods can help inform each other. Particularly in light of the time constraints OIOS-IED evaluation teams face in their data collection, it is important to plan ahead and determine whether it is most desirable to conduct focus groups before, after or in parallel to interviews. Sometimes, for example, it is desirable to save focus groups until the end of a short field mission and, armed with a better understanding of the key issues as a result of interviews, invite select interviewees to attend focus groups sharply focused on a narrow set of very specific issues; [IED #34].

- As in all aspects of OIOS-IED's work, in planning specific focus groups it is important to ensure maximum geographic and gender balance in forming focus groups;

- Focus groups are meant to capitalize on group dynamics, but the multicultural nature of the UN can make these dynamics particularly challenging to manage - e.g., gender, culture and hierarchy might result in uneven patterns of participation and non-participation levels, with some participants dominating and others remaining silent. This can be managed when it arises in the group, but to the extent the evaluation team can foresee and plan groups accordingly in advance, this challenge can often be pre-empted;

- Issues of language can exacerbate these challenges. In some cases, a participant will feel intimidated from participating because of a lack of confidence in his or her proficiency in the language being used in the focus group. If this is a foreseen concern, the evaluation team might consider separating the focus groups by language;

- Similarly, any known information about specific inter-personal dynamics can help the evaluation team plan a successful focus group. For example, knowledge of inter-divisional rivalries might preclude two individuals being included in the same group, or from seating them next to each other in the group; and

- If being held during a breakfast or lunch hour, and if resources permit, consider providing light refreshments. Small gestures such as this can help encourage attendance.

Conducting focus groups

The data collection tool for conducting focus group discussions is called a focus group guide. OIOS-IED has a number of sample focus group guides on which to draw [IED #34].

All of the guidance pertaining to the development of interview guides also applies to focus group guides. That said, a few specific points are worth bearing in mind when developing focus group guides:

- It is important to include a point underlining the purpose of the focus group - and the desirability of differences of opinion, where these exist - so as to solicit a productive, constructive exchange of views;

- Develop a funnel-shaped series of no more than eight questions, starting with a few general questions and moving to two or three key questions of greatest importance; and

- For those questions on which the evaluation team foresees a particularly rich exchange of perspectives, it is often useful to include specific probes to ensure that these potential differences are actively explored - e.g., by including a probe such as the following: "What do others think?" or "What if any other perspectives on this issue do others have?".

The role of the evaluation team member conducting the focus group is that of a focus group moderator. The role of the moderator is very important in focus group sessions. The moderator's main tasks are to:

- Ensure that all of the critical questions in the focus group guide are covered;

- Facilitate a natural conversation among the participants themselves;

- Enable any salient points that were not originally foreseen in the focus group guide to surface and be reflected upon by the group;

- Allow the natural conversation described in the previous points to occur, while at the same time preventing the conversation from straying from the central topic and knowing when a given line of questioning has been exhausted and it is time to move on; and

- Manage the group dynamics to ensure that all voices are heard, and that the similarities and differences of perspective are explored, without allowing any individual to dominate the conversation, or allowing differences of opinion to escalate into conflict.

The main pointers worth bearing in mind when moderating focus groups are the following:

- The moderator should convey a sincere welcome to all participants as they arrive. This puts participants at ease and is more likely to elicit a frank and open exchange of views during the session;

- In addition to covering the same introductory points as those used for interviews, the moderator should also state the general ground rules for discussion at the outset - e.g., those related to equal participation, the desirability of frank but respect-

ful debate, including disagreement where this is present, the potential need for the moderator to interrupt the conversation occasionally to steer the conversation back to the topic;

- The principle of active listening is even more challenging in focus groups than in interviews, as the moderator's focus is on facilitating a focused conversation while comparing and contrasting the various perspectives being provided. For this reason, wherever possible, an OIOS-IED note-taker should be present in the focus group. After each session, note-takers review and write additional notes where appropriate;

- Before embarking on the focus group questions, the moderator should ask all participants to undertake a tour-de-table to introduce themselves. The moderator might consider using name tags, sign-in sheet, and/or physical map of participants' location around the table, in order to keep track of names and, where appropriate, refer to them by name. This personal touch can contribute to building positive rapport - and help the team keep track of focus group participation numbers and any necessary follow-up;

- The moderator should also consider using easels, whiteboards, chalkboards or laptops to summarize the feedback received during the group discussion and to help the group explore areas of convergence and divergence of perspectives;

- The main challenge of focus groups is managing the conversation so that they are free-flowing yet focused, and the group dynamics so that they are frank, open and inclusive yet congenial. All of these goals must be achieved within a limited time. As a result, the moderator must sometime interrupt and redirect the conversation back to the topic, ask a dominating participant to withhold his or her views and allow less vocal participants to speak, and undertake other measures to manage the group. It is important to manage the focus group in a professional, positive manner; and

- Often, the first few focus group sessions yield considerable new information. As more and more sessions are conducted, information can sometimes start to become repetitive. This is referred to as "saturation."

SELF-ADMINISTERED SURVEYS

What they are and how OIOS-IED uses them

Through self-administered surveys, OIOS-IED frequently collects information from a large number of respondents on a series of distinct questions or issues (Box 5). As such, surveys are a potentially powerful tool for establishing the magnitude of a given phenomenon or sentiment that can be generalized to the larger survey population.

There are many external resources available that summarize good practice in the use of surveys [➔ World Bank. 2009a] [➔ Shannon et al. 2012a] [➔ Dillman et al., 2008a] [📖 BetterEvaluation].

Surveys typically obtain data on various types of information, including: background, descriptive data, behaviours, attitudes and beliefs, opinions, satisfaction and knowledge. Toward this end, surveys ask individuals both closed-ended and open-ended questions. Closed-ended questions offer a limited range of responses and so produce information that can be quantitatively summarized - and, where necessary, disaggregated on key variables such as gender, geographical location or staff level. Open-ended questions produce qualitative data that, if and when aggregated through the process of data coding, can provide a valuable quantitative overview of the magnitude of respondents' views.With self-administered surveys, control of the data collection instrument, or questionnaire, rests with the respondent. The opportunity for clarifying or explaining questions is no longer available. It is therefore crucial that the evaluation team ensure that there is no ambiguity in the survey questionnaire - and that they pretest questionnaires before deploying them more widely. Poorly designed questionnaires not only cost respondents' valuable time; they also limit the validity and reliability (Part II,

Section 1.1) of the data they contain and undermine OIOS-IED's credibility.

Surveys are sometimes criticized as being too limited in their benefit to the evaluation to warrant the considerable effort they entail. However, when executed properly, they often represent the one and only source of statistically valid and reliable source of information on stakeholders' views and experience.

The main strengths of self-administered surveys include:

- They are especially useful when broad information from a large population is needed;
- They collect comparable data in a systematic manner;
- They are generally less costly than interviews;
- Evaluators can ask relatively complex questions, since surveys allow time for respondents to reflect on events and report changes and feelings;
- They collect systematic and comparable data using standardized measurement;
- Responses to closed-ended questions can be easily summarized (usually done electronically); and
- They allow for anonymity of responses, which might encourage subjects to answer sensitive or embarrassing questions.

The main limitations of self-administered surveys include:

- The amount of time required for designing, pre-testing and conducting them;
- They can be subject to survey error;
- The evaluator cannot control who takes the survey, or when and why;
- There is usually no opportunity to clarify responses if these are not clear;
- Response rates can be low, rendering follow-up a labour-intensive endeavour and making it difficult to generalize results for the survey population; and

- They are inflexible in that they require the initial design to remain unchanged throughout the data collection.

Box 5: A Common Use of Self-administered Surveys in OIOS-IED

OIOS-IED often uses self-administered surveys to collect data to take stock of relevant aspects of an evaluand's operations. In this type of surveys, OIOS-IED sends the survey questionnaire to individual entities, typically to the evaluand focal point, to survey particular practices, approaches, activities or other elements of the entity. The focal point may complete the questionnaire on behalf of his/her entity, or may delegate the task to one or more officers in the entity. This type of survey is particularly useful in thematic evaluations as well as the Biennial Reports and Scorecard Reports, as they allow direct, objective comparisons across entities. They can also be used in programme evaluations, in order to survey subprogrammes (i.e., divisions, units, or other substructures) within the programme, regional offices, country offices (or, in the case of peacekeeping evaluations, peacekeeping operations) selected for case study.

Two main considerations should be taken into account when conducting this type of survey:

- Evaluand focal points are often extremely busy and their internal colleagues might require significant time to gather the information being asked, so it is important to (a) limit the number of items asked to those that OIOS-IED actually intends to use; and (b) allow ample time for the focal point to collect the information requested; and

- For optimal validity and reliability, the self-reported information obtained from these questionnaires should be accompanied by a request for, and examination of, supporting documentation (also known as "means of verification").

Planning self-administered surveys

When planning a survey, the following points should be considered:

- Determine early on whether it will be most beneficial to the evaluation to deploy the survey before, after or in parallel with other data collection methods;
- Determine as early as possible whether the survey should appear in multiple languages, as this can have implications for the evaluation budget and/or

the timing of the survey's pre-testing and deployment; and
- Determine any protocols that must be followed (Box 6).

In order to help hedge against low response rates, at the planning stage:

– Alert stakeholders of the survey early on and underline the importance of their participation;

– Consider requesting the evaluand's USG or other senior-level manager to send an email just prior to the survey's deployment, alerting staff to its imminent deployment and underlining the importance of their participation; and

– Determine the time frame that will make most strategic sense. Surveys are less likely to yield acceptable response rates during major holiday periods, key evaluand reporting or strategic planning periods, and major conferences or other events critical to the evaluand.

Box 6: Surveying Member State Representatives

Planning surveys of Member State representatives requires particular care and should be considered carefully. Typically, the protocol followed for these surveys involves correspondence from the OIOS USG to the Mission Head, with the survey being administered in paper and/or electronically [IED #35 - 38]. Protocols to follow for representatives of other inter-governmental bodies should be determined in consultation with the secretariats of those bodies.

Designing self-administered surveys

All survey questionnaires should be easy to understand and navigate for the survey respondent, and should yield unambiguous data that are analysable and comparable

for the OIOS-IED evaluation team. They should also be as short as possible, focusing on the most critical items, rather than what is merely interesting but not essential to the evaluation at hand.

Beyond this broad level, evaluation teams should consider the following points for designing high-quality questionnaires:

- Begin with an introductory page that briefly spells out the objective and purpose of the survey (and the OIOS-IED evaluation), gives respondents an accurate gauge of the approximate survey duration, and states OIOS-IED's confidentiality guarantee;

- Structure the survey questionnaire in a logical manner that respondents will find easy to navigate. Order questions around similar topics and in a logical way, beginning with an appropriate first question and placing potentially objectionable questions at the end;

- Use clear and simple language that avoids jargon;

- Ask questions as complete sentences;

- Define key terms that are not known to be widely familiar to respondents;

- If necessary, arrange for translation, most typically into French and/or Spanish;

- Avoid loaded questions - i.e., ask only one question at a time, with only one thought or idea per question;

- Avoid leading questions - i.e., frame questions in a neutral manner, free of any language or format that biases the respondent to answer in a certain way;

- Provide specific time references when asking about past events; and

- Thank respondents for their time at the beginning and the end.

One of the main choices when designing a questionnaire is the type of questions to use. There are two basic types of questions:

- Closed-ended, or structured, questions that offer a fixed set of responses and result in quantitative data; and

- Open-ended, or unstructured, questions that do not offer any response categories and result in qualitative data (that later can be quantified through coding).

A combination of both types of questions is typically used, although survey research has recommended that the number of open-ended questions be kept to a minimum since they tend to negatively impact on response rates.

A number of specific points are worth bearing in mind in developing closed-ended questions:

- Determine which types of closed-ended questions of Table 14 are most fitting, in light of the information sought from the question at hand;

Table 14: Types of Closed-Ended Questions

Type of Closed-ended question	Type of Response	Examples
Nominal	Unordered response categories	A "check all that apply" list of mutually-exclusive response categories.
Ordinal	Ordered response categories	A rating scale such as "strongly disagree", "disagree", "neither agree nor disagree", "agree" and "strongly agree." It is recommended that, when using this type of response set, there be either five or seven response categories with a neutral mid-point. This feature gives the range needed to detect differences and provides an opportunity for those with a neutral opinion to express that point of view.
Numerical	Numerical responses	"How many years have you been working in the UN?"

- Obtain precise estimates where possible;

- Use mutually exclusive response categories;

- Use balanced scales with an equal number of positive and negative responses;

- Label each point of the response scale (or at least the end and mid points);

- Consistently position all scales to read from most positive to most negative, or vice versa. Do not confuse respondents by switching scale directions, as this can lengthen the time it takes to complete the survey as well as increase response bias;

- Wherever identical information is being sought across a number of items, repeat the exact same scale rather than creating a similar but different one - e.g., do not switch from a four-point agreement scale to a five-point agreement scale unless absolutely necessary;

- Consider using a response matrix (e.g., "How strongly do you agree or disagree with each of the following statements?"), rather than asking individual questions sequentially using the same scale, as this can lead to respondent fatigue;

- When using response matrices, consider rotating or randomizing items in the inventory of questions so as to avoid response bias;

- Consider the advantages and disadvantages of permitting such response categories as "Don't Know," "No Opinion," "Not Applicable" or "No Basis for Judgment"; and

- Consider the advantages and disadvantages of forcing certain questions to be mandatory for respondents and which are to be voluntary.

A few points are worth bearing in mind in developing open-ended questions:

- It is a good rule of thumb to try to limit the number of open-ended questions to one to three. Restricting the number of open-ended questions also minimizes the risk of collecting data that are not used, and enhances response rates;

- Open-ended does not mean imprecise. Better information is obtained when the open-ended question is as specific as possible. Questions framed simply as "Additional comments?" and "What else?" often yield low response rates, vague responses that are very difficult to code, or both; and

- To boost the richness and analyzability of the responses provided, consid-

er giving respondents brief instructions such as, "Please be as specific as possible" or "Please provide 1-2 specific examples" or both.

As with interview guides and focus group guides, survey questionnaires should always be pre-tested before they are finalized, first by evaluation team members and then with more than one prospective respondents. The pre-test should estimate the time to complete the survey, identify overall logical flow and clarity of language, test skip patterns and mandatory questions, obtain feedback on the clarity and order of the questions and the response categories, and obtain feedback on the ease of using the instrument.

Conducting self-administered surveys

The data collection tool for conducting self-administered surveys is called survey questionnaire. OIOS-IED has a number of sample questionnaires on hand to draw on in developing one's own survey instruments [IED #39 - 46].

OIOS-IED surveys are typically conducted by web using an online survey platform. On rare occasions, OIOS-IED uses paper-and-pencil surveys.

During the survey period, which typically lasts three to six weeks, OIOS-IED evaluation teams continuously monitor response rates. The response rate calculation is simply the ratio of the total number of completed valid survey questionnaires received (in the numerator) to the total number of respondents the surveys were sent whose email address was deemed valid (in the denominator).

Ways to enhance response rates include:

- Conducting the survey at an optimal time;
- Limiting the length of the survey questionnaire;
- Targeting respondents' awareness of the survey and its closing date;
- Using a minimum of two politely-worded reminders to non-responders;

- Working out any technical issues such as connectivity and programming "bugs" (in web-based surveys);
- Convincing members of the survey population that their responses will be used and will add value to the evaluation;
- Building the trust of survey respondents with regard to the confidential handling of their responses; and
- Designing the questionnaire so it is easy to navigate and understand the questions.

When conducting any survey, it is important to consider the four types of possible survey error - i.e., those factors that reduce the quality of survey data:

- Sampling error (also margin of error) - the result of the fact that the survey was conducted among one particular sample of the universe;
- Measurement error - the result of imperfect data collection such as due to poor question wording or flawed rating scales;
- Coverage error - the result of not all units of the population having an equal chance of being sampled for participation in the survey; and
- Non-response error - the result of people who respond to a survey being different from sampled individuals who did not respond, in a way relevant to the study. To assess non-response error, OIOS-IED evaluation teams undertake a non-respondent analysis or imputation.

Data from surveys with low response rates should be used with caution. In these cases, it is not possible to draw valid inferences to the general population since the sampling error (margin of error) is too high. OIOS-IED evaluation teams state survey response rates clearly. They treat low-response-rate data as indicative but not representative.

FIELD-BASED SURVEYS

What they are and how OIOS-IED uses them

OIOS-IED also uses field-based surveys (also called "local population surveys") in order to obtain information from programme beneficiaries. These surveys are particularly useful in measuring the effects and/or impacts of a given programme on the population(s) whom the programme is intended to benefit.

There are a number of external resources that summarize good practice for field-based surveys [➔ Tourangeau. 2014a] [➔ Ford. 2009a] [➔ Cohen. 2011a]. For them to succeed, it is important to pay close attention to sampling strategies and survey administration protocols.

From a human rights perspective, field-based surveys can also reveal whether the UN is reaching the populations it should be - a core aspect of the relevance criterion in evaluation. It is therefore important that OIOS-IED's field-based surveys not only attempt to include known beneficiaries of the programme, but also those who might not have benefitted but were otherwise eligible for services. Thus, from a human rights-based perspective, field-based surveys can pinpoint the reasons why some qualified individuals are not benefitting from the programme's work - e.g., lack of information, misperceptions, perceptions based on real shortcomings in the programme, barriers to access, and so on.

In addition, it is worth noting that information yielded from field-based surveys is not limited only to the opinions of programme beneficiaries. These surveys can also serve as opportunities to collect objective information on the programme's effectiveness through direct observation, such as what the purported beneficiaries of food assistance in humanitarian situations have in their cupboards.

The main strengths of field-based surveys are that they:

- Generate up-to-date primary data on the direct and indirect impacts of the programme and the extent to which the programme is making a difference in the lives of the beneficiary population(s);
- Convey to beneficiaries that their voice matters in shaping a programme's direction; and
- Collect evidence among low-literacy beneficiaries (when conducted using in-person interviewing).

The main limitations of field-based surveys include:

- Their time consuming and potentially costly nature if they are to achieve an adequate sample size for extrapolating survey data to the wider population;
- Unclear sampling strategies (particularly in lesser developed countries or territories), where census bureaux are either non-existent or census figures, upon which sampling is typically based, are unreliable for a variety of reasons;
- Potentially difficult accessibility to beneficiaries, coupled with logistical and security challenges of getting to the most difficult-to-reach areas;
- Cultural challenges (such as sensitivities around the subject matter being discussed and reluctance for female respondents to engage with male interviewers);
- False expectations among beneficiaries that the survey researchers will solve their disputes with the UN or others; and
- Quality and reliability of data cannot be directly controlled if outsourced to a third party.

Planning field-based surveys

When planning field based surveys, the following questions must be addressed:

- What type of data will be obtained - e.g., perceptual data only, or some other type?

- From whom will data be obtained (i.e., the sampling frame) - e.g., beneficiaries of a programme, or non-beneficiaries as well; people affected by a humanitarian

situation or those not affected but living in nearby communities as well; young and old alike?

- How will data be obtained - e.g., through real-time tablet-based data entry, self-administered paper-and-pencil format, or enumerator-administered format?

- Who will obtain the data - e.g., an international or locally-based private sector research firm, an international or locally-based academic institution, one or more individual consultants, or some combination of these?

- How much will it cost to obtain the data - which is largely dependent upon the size of the population at hand, and the corresponding sample size needed to obtain an acceptable level of confidence and margin of error?

OIOS-IED typically hires consultants to conduct its field-based surveys, so adequate time must be built into the project timeframe to identify the consultants, issue their contracts, and train them. In addition, owing to their unique complexities, field-based surveys can require on-going management and trouble-shooting on OIOS-IED's part. This additional management challenge can place significant additional time and re-source demands on those managing the survey prong of data collection.

Some points from OIOS-IED's previous experience should be considered when planning field-based surveys:

- Many OIOS-IED field-based surveys are of local populations, and it is important to explore how the universe of that population will be identified, for example through census data or some other means. When sampling, it is important to rely on a reputable sampling frame. Oftentimes this consists of a national census, but even here the assumptions of the census sampling frame must be examined. In addition, in conflict-affected and post-conflict contexts, the basic assumptions

of representative sampling are absent - population displacement as refugees or internally displaced people, genocide of a certain group, and so on, can affect the fundamental tenets of sampling;

- It is also usually advisable to contract local enumerators to assist in the data collection effort - and to train them in advance in order to ensure maximum fidelity to the survey instrument;

- Surveys must be translated into, and pre-tested in, the local language(s). This can add significant cost;

- It is important to bear in mind the ethical as well as the methodological considerations associated with field-based surveys. Consuming respondents' time for an hour or more can detract from their livelihoods. It is therefore important to explain clearly at the outset of each and every interview the nature of the interview and what can and cannot be expected.

Finally, the very notion of engaging directly and systematically with local populations can be a politically challenging undertaking in many contexts. This can have practical implications for OIOS-IED, in that it must often lay the groundwork with national and local officials, obtain permissions where necessary, make the necessary overtures to local community leaders, and so on, before embarking on the survey itself.

Conducting field-based surveys

The data collection tool for conducting field-based surveys is called field-based survey guide. OIOS-IED has sample survey instruments to draw on in developing one's own field-based survey guides [IED #47 - 49].

In developing field-based surveys, most of the same design guidelines that apply to self-administered surveys are relevant. Other noteworthy tips include the following:

- Oftentimes, visual aides, such as a smiley face or frown for a satisfaction response scale, can produce more valid and reliable data than numeric scales;

- Because of malnutrition, disease, post-traumatic stress and other factors present in many corners of the world, it is sometimes helpful to include a brief set of items to test potential respondents' cognitive ability before proceeding to the survey questions;

- Points about ethical aspects of survey administration are particularly salient in field-based surveys. Vulnerable populations are often particularly sensitive to questions being asked, so it is important that questions be framed to avoid causing distress as much as possible. Enumerators must clearly communicate to respondents that they are free to skip any items they do not want to answer, and exit or postpone the interview at any time should they so choose;

- At the same time, owing to the significant disincentives to participation in OIOS-IED's surveys, it is oftentimes necessary to equip enumerators with a series of instruction scripts to counter initial refusals to participation and to know when "no" finally means "no"; and

- The national language is often not universally spoken or read in every area of a given country. Translation into one or more additional local languages might therefore be necessary.

Since field-based surveys are administered by enumerators, it is critical that the enumerators be properly trained and that quality control measures are put in place to ensure that the survey is being administered consistently for maximum fidelity to the survey instrument. This would include training on the survey and direct observation of the enumerators in the field (announced and unannounced) to provide feedback on survey administration. It is standard practice to request regular updates from enumerators to track response rates, identify any potential barriers to participation or comple-tion, and identify solutions. This is also an opportunity to monitor enumerator safety and security issues.

DIRECT OBSERVATION

What it is and how OIOS-IED uses it
A method that OIOS-IED increasingly uses is direct observation. As implied by the term, it is a process by which data are generated through the direct observation of a situation, group or event.

Direct observation is often identified as a qualitative data collection method. However, when used with a structured observational guide, it can also produce quantitative data. Because of the time and resource constraints under which OIOS-IED operates, the use of direct observation must be selective, looking at a few activities, events or phenomena that are central to answering the evaluation questions.

There are a number of external resources that summarize good practice for direct observation [➔ World Bank. 2009a] [📖 BetterEvaluation].

OIOS-IED's use of this particular data collection method in the past has included observing inter-governmental processes, programme training and/or outreach pro-grammes, and the delivery of operational activities in the field.

Advantages of direct observation are that:

- It relies less on the perceptions of respondents and more on the real situation being observed, including actual behaviours and body language;
- It enables OIOS-IED to study an event, an institution, a facility or a process in its natural setting;
- The observer can consider the context in which the observation occurs and not just the condition or behaviour;

- The observer can be more open to discovering new and different perspectives;
- The observer can obtain from the observation information that people would be reluctant to discuss (sensitive issues); and
- It avoids issues associated with the passage of time, such as memory decay when using interviews or questionnaires.

Limitations of direct observation include:

- It can be labour-intensive and costly;
- Unless sites are carefully sampled, there might be bias in site selection;
- The knowledge that OIOS-IED is observing one's activities can change the behaviours of those being observed; and
- It can be difficult to identify or exclude observer bias.

Planning direct observation

When planning direct observation, OIOS-IED evaluators must determine the number and type of direct observations required for maximum confidence and precision. Additionally, if individuals are involved, a decision must be taken on whether to observe and record their behaviour with or without their knowledge. Observation with subject knowledge is more common, but poses a potential problem - individuals might act or respond differently than if there was no observation, which in turn can lead to false interpretations and problems with bias in the data. However, observation without their knowledge raises a question of ethics in evaluation research.

The following points should also be kept in mind when planning direct observation:

- During the evaluation design stage, it is important to try to get an early sense for what specific observational data might help answer the evaluation questions;

- It is also helpful to identify observational opportunities elsewhere at UN Headquarters that are of specific relevance to the evaluation; and

- It is sometimes necessary to request official permission in advance to attend a specific event or visit a specific site, either with senior management, the local authorities, or both.

Conducting direct observation

The data collection tool for conducting direct observations is called a direct observation instrument. OIOS-IED has sample direct observation instruments in place to draw on in developing one's own instruments [IED #50 - 55].

As with interviews guides, the OIOS-IED evaluation team must choose whether to use a structured, semi-structured or unstructured instrument to record observations. When it is clear what the evaluator is looking for, a structured or semi-structured observational guide that allows for recording standardized and systematic information is appropriate. By virtue of the forethought that goes into selecting events or sites for observation in the first place, most of the observations OIOS-IED evaluation teams undertake entail structured or semi-structured guides. Common exceptions are when a more exploratory approach is required in the evaluation, or when an evaluation team member is invited to a meeting with little advance notice, and there is little time to think through what s/he wishes to observe.

A structured observation guide is akin to a survey questionnaire that includes both closed- and open-ended questions to be answered through one's own observations. A semi-structured observation guide is closer to an interview guide. It consists largely of open-ended questions, in response to which the observer records his or her observations. Closed-ended questions might include such items as the number of times someone comments on the evaluand during a meeting, how many of these comments were positive or neutral, whether the meeting begins and ends on time, and so on. Open-ended questions might be used to describe body language of attendees, the overall condition of a visited site, overall observations of group dynamics, and more.

In addition to written observations, OIOS-IED evaluators can also consider the desira-

bility and appropriateness of using supplementary means of recording observations, such as photographs.

As with field-based surveys, when using a structured or semi-structured observation instrument, it is necessary to train all evaluation team members in its use in order to ensure fidelity to the instrument. Without a consistent application, reliability of the data might be questioned. It will not be known if the variation recorded comes from differences in the actual situation observed or in the different applications of the guide. Wherever possible, it is also ideal to have two team members conduct the observation, in order to minimize observer bias.

Although each direct observation is unique, a number of points of guidance can help make direct observations as profitable as possible. These include the following:

- In direct observations, both active listening and active observing are key. Being alert to social interactions - who is speaking more and who is speaking less, how individuals react to each other's comments, overall body language, what is going on in the surrounding environment, and even metrics such as temperature or noise, are at least as important as what is being said in these instances;

- It is important to avoid the Hawthorne effect, also known as the desirability effect, as much as possible when undertaking observations. The Hawthorne effect is simply the dynamic that often occurs when stakeholders know evaluators are observing them - they can alternatively be on their best behaviour and avoid anything that might reflect negatively on them or their work, or they might use the event as a platform to convey their opinions more loudly or set a group dynamic in motion for the evaluators to observe. For this reason it is preferable, wherever possible, to keep meeting attendance as impromptu wherever possible - and even avoid being overly obtrusive as the evaluators. The observers should at least position themselves in a spot that is least disruptive to the event being observed. Ideally, they would not even be noticed by the event participants;

- It is advisable to clean the notes from the observation immediately afterward, while memory is still fresh. It might also be helpful, when appropriate, to ask follow-up questions to the evaluand in order to better understand the data obtained from the observation.

DESK REVIEW

What it is and how OIOS-IED uses it

OIOS-IED evaluators collect a broad range of documentation on the evaluation topic from both primary and secondary document sources. These include legislative documents, policies, strategies, self-evaluations, IMDIS (Integrated Monitoring and Documentation Information System) data, financial records, work plans, project documents and photographs, among others. OIOS-IED uses systematic desk review (also known as document review) to collect rich qualitative and quantitative data on the programme it is evaluating.

There are a number of external resources that summarize good practice for desk reviews [📖 BetterEvaluation].

Systematic desk review entails the structured review of key documents - whether to look for specific data points or to cull a sample of documents from a much larger universe of their kind - using a data collection tool to answer specific evaluation questions across a series of documents of the same type. One way to think about systematic desk review is that it is a means of surveying. Rather than asking questions of respondents, however, the evaluator is asking questions of documents.

The data garnered from desk review can serve as a powerful source of evidence to be triangulated against the other data sources gathered during the evaluation. Examples of desk reviews that can inform the overall analysis, broken down by evaluation criterion, include the following:

- Effectiveness/Impact - a systematic synthesis of existing evaluations, reviews, after-action reviews, lessons-learned and other evaluative reviews, to plot out those targeted results the programme was found to be effective in achieving, those it was found less effective in achieving, and why;

- Efficiency - a systematic analysis of time-to-deployment data from human resource records to determine response time in an emergency or costs of a programme or a project compared to its known (or unknown) results to obtain a proxy gauge of cost-effectiveness; and

- Relevance - a systematic review of project documents in a specific country against the programme's Strategic Framework, the UN Development Assistance Framework (UNDAF), and needs assessments, in order to determine the overall alignment of the project to the programme's mandate and to local priorities.

Planning desk reviews
After determining what data points should be gathered through desk review, OIOS-IED evaluation teams must next determine the type and number of documents required for maximum confidence and precision. Give the time and resource constraints inherent in OIOS-IED evaluations, it might be necessary to draw a sample (Part II, Section 1.1) of documents to review, rather than reviewing the universe.

Conducting desk reviews
The data collection tool for conducting desk reviews are typically called desk review instruments. These instruments should list clear criteria for assessing each element of the documents being reviewed. Similar to direct observation, a structured instrument to record the content of the documents reviewed will capture standardized and systematic information.

OIOS-IED has samples of desk review instruments in place to draw on in developing one's own [IED #56 - 57].

One way of thinking about the desk review instrument is as a survey of documents - instead of asking questions of a respondent, the evaluator is asking questions of a document. Common questions include the following:

- How often does each document mention a term being sought?
- How clearly does the document explain its rationale and logic?
- What does the document say about the programme's effectiveness, efficiency and relevance?
- When was the document's first draft released and when was it finalized?
- Who was and was not included in the process of developing the document?

When using a structured desk review instrument, it is necessary to train all concerned evaluation team members in its use in order to ensure maximum fidelity to the instrument. Without consistent application, reliability of the data might be questioned. Information from a desk review gathered in a structured manner can be represented numerically.

The following guidelines should be followed when conducting desk reviews:

- All desk review instruments should be piloted, discussed within the evaluation team, and refined early in the review;

- All team members should be trained in how to use the instruments to increase consistency and therefore reliability of the review;

- OIOS-IED is typically not a technical expert in the substantive area covered by the document, so it is important to recognize the limits of the team's ability to assess the documents at hand;

- Wherever possible, two or more team members should form a dyad or triad and read the same document, in order to avoid reviewer bias;

- The team leader should monitor fidelity to the data collection instrument after the first few documents reviewed, and reconcile differing interpretations (and where necessary correct misinterpretations) of what is being asked in the data collection instrument or the criteria it is using; and

- At the end of the desk review, consider calculating an inter-rater reliability score to validate the fact that a high degree of consistency is present across the members of a dyad or triad. [📖 NCBI] [📖 MED-ED].

CASE STUDIES

What they are and how OIOS-IED uses them

Generally, case studies attempt to learn about and understand a complex issue through an extensive description and analysis of that issue as represented by one or more particular "cases" or units, in their entirety. Case studies apply the data collection methods described above - interviews, surveys, direct observation and desk review - in order to obtain in-depth and comprehensive information on the case(s) examined. Although the unit of analysis in case studies is typically countries, other units, such as events, individuals and outputs are also good candidates for case study.

There are a number of external resources that summarize good practice for case studies [➔ World Bank. 2009a] [📖 BetterEvaluation].

To ensure meaningful information, OIOS-IED evaluation teams typically sample case studies - i.e., select a sub-set of all possible case studies -- in a systematic manner. Teams also agree on the type of case study to use. Examples include:

- Illustrative case studies - these are mainly descriptive studies that attempt to portray the programme in-depth, and as realistically as possible, within its policy context;

- Exploratory case studies - while this is mainly descriptive, its goal is to generate hypotheses about the programme that can later be tested, using quantitative of qualitative methods;

- Critical instance case studies - this singles out a specific and often unique case in order to investigate its problems and strengths. It attempts to learn from the uniqueness of the programme;

- Programme implementation case studies - as the name implies, this is an investigation of how the programme has been implemented and is operating. It usually includes a number of programme sites;

- Programme effects case studies - here the focus of interest shifts to the end results of the programme and attempts to deal, qualitatively, with the question of causality. It, too, usually involves several programme sites; and

- Cumulative case studies - this approach utilizes evidence from several programme sites to answer a full range of evaluation questions.

The specific sampling method for selecting case studies will depend on which of the above types the evaluation team wishes to undertake. Such sampling methods can range from simple comparison of all potential units on a small number of cases (which is especially useful in critical instance case studies, and in instances of small universes, such as PKMs or SPMs) to more elaborate comparison using cluster analysis (which is particularly useful in the other types of case studies, and when sampling from among a large number of cases).

The main advantages of case studies are that they:

- Have the ability to develop the needed information with a relatively small number of cases;

- Can provide the information on general trends across cases that can be used to assess how a programme is or has worked;

- Allow the evaluator to experience "real" programme examples in their entirety, which can give added insight for the evaluation; and

- Are a highly flexible approach that can be applied in many situations and often when other approaches are impractical.

The main disadvantages of case studies are:

- There is increased opportunity for bias to be introduced into the results, because of the extent of comprehensive involvement in the cases;

- The approach might not be as analyzable if not undertaken systematically with a clear plan for comparing, contrasting and aggregating disparate case studies chosen ; and

- The heavy focus on context makes it difficult to generalize the results to the larger universe of programmes.

Planning and conducting case studies

As with any other data collection approach, using a case study should involve a specific plan for how to proceed. When conducting a case study, the following steps are recommended:

- Develop the objectives of the case study;

- Develop the specific questions that will be answered with the data collected using the case study;

- Select the case study approach and the specific cases to be included; and

- Determine the data collection methods that will be used and how the data collected will be analyzed.

OIOS-IED has sample case study summaries to draw on in developing one's own case studies [IED #58].

FIELD MISSIONS

What they are and how OIOS-IED uses them

As a rule of thumb, OIOS-IED tries to collect as much data as possible through remote means, in light of its resource and time constraints and its desire to be as cost-conscious and have as low a carbon footprint as possible. That said, field missions are often indispensable to the credibility and utilization OIOS-IED seeks in its evaluations.

The major advantages of field missions are that they:

- Provide an opportunity to interview stakeholders who would otherwise be difficult or impossible to reach by remote data collection means - e.g., ministerial officials, beneficiaries, community-based organizations;

- Constitute the only way in which OIOS-IED can undertake direct observation of programmes in their day-to-day work; and

- Contribute to utilization of the evaluation by continuing to generate interest in and demand for the evaluation by key stakeholders, and to the credibility of the evaluation by underlining to the evaluand that OIOS-IED has literally seen what it does and can therefore understand it better.

The major disadvantages of field missions are that they:

- Are expensive and time-consuming to conduct, and leave a carbon footprint;

- Cannot be used to generate data that can be generalized to all locations or countries; and

- Can introduce bias into the evaluation.

Planning field missions

Field missions involve the use of interviews, focus groups, direct observation and/or document review. In planning for a field visit, the following steps are taken:

- Coordinate with evaluand focal point regarding the timing and logistics of the mission;

- Based on stakeholder analysis, identify groups and individuals for interviews/focus group discussions during the mission;

- Sample sites and/or events for direct observation;

- Identify any documents to collect and/or review while on site;

- Work with the evaluand focal point to establish a meetings schedule, including an entry meeting with the mission focal point upon arrival to go over evaluation objective, mission programme and any logistical or security matters; and

- Prepare for interviews, focus group discussions and direct observation.

The planning of field missions is an area in which OIOS-IED must often be most vigilant in balancing its mandated independence with its commitment to evaluand consulta-

tion. The inputs and advice of the evaluand focal point and others are vital to a solid mission itinerary. However, OIOS-IED must also independently identify (through an independently drawn sample, if necessary) those stakeholders it feels a need to meet, as well as those direct observations it wishes to conduct. Given a large number of countries in which some programmes operate, it is often necessary to strategically select those that will be included for field missions and those that will be excluded.

Most broadly, it is important to take a strategic, risk-based approach to the selection process. Important considerations when selecting countries or sites might include the following:

- What is the overall financial risk profile of each country?

- How long has the programme been present in each country - i.e., such that longer-term impacts might be explored in addition to shorter-term effectiveness?

- Is the programme still active in the country - i.e., such that interviews with staff and other stakeholders will be relatively straightforward to arrange, and direct observations will be possible in addition to interviews?

- How much of the programme's programmatic profile is covered in the country - e.g., is just one subprogramme's substantive areas of focus represented or are multiple substantive areas covered, are both its normative and operational pillars covered or just one, does it entail both emergency and non-emergency interventions or just one, and so on?

- What if any evaluations have already been conducted of the programme's projects or overall operations in each country?

- What if any important typological variables are relevant in the evaluation that speak to specific attributes of each country - e.g., peacekeeping operation (PKO)

versus special political mission (SPM) versus other, emergency context versus non-emergency context, Delivering as One country versus non-Delivering as One context, overall youth population, and so on.

There are any numbers of methods to help OIOS-IED evaluation teams select countries for inclusion as country case studies in an even more systematic fashion. A particularly powerful method, increasingly used in country case study selection, is cluster analysis. Cluster analysis is an exploratory data analysis method that enables evaluation teams to sift through large numbers of variables on a large number of cases (such as countries) to detect patterns among these cases that are difficult for the naked eye to detect. Cluster analysis reveals underlying typologies of a programme's work at country level, enabling a more nuanced comparison of its global operations. External resources help explain this method [➜ Hair et al. 2009a], and OIOS-IED has developed its own instruction sheet for conducting cluster analysis in SPSS Statistics [IED #59].

There are a number of external resources that summarize a case study approach to field missions [➜ Hancock et al. 2011a] [➜ Stake. 1995a] [➜ Yin. 2011a] [➜ Yin. 2013a].

Conducting field missions
Typically, there are no specific data collection instruments geared specifically for conducting field missions as a method in itself, other than those to be used for interviews, surveys, direct observations, and desk reviews to be conducted in the countries or sites selected.

That said, there are any number of tools that might assist OIOS-IED evaluation teams in ensuring that the various data sources that inform the analysis of the country case study are compiled systematically, and preferably in "real time" as they are gathered and/or analyzed. Simple tools such as spreadsheets can help collate this information across all of the countries selected. This central organizing tool is important because it is frequently the case that evaluation team members undertake disparate strands of the data collection effort. Spreadsheets summarizing the data gathered on each

country help ensure that the full team has a holistic overview of all data on all countries included in the analysis.

Upon return from a field mission, OIOS-IED evaluation team members typically complete mission reports [IED #60], aide mémoires or mission briefs to document the work conducted. These post-mission documents are important for accountability purposes and are critically important to ensure systematic cross-comparison across field missions (especially in cases where evaluation team members split up and cover different locations), and to ensure the reliability of the data collected - i.e., recording one's observations as soon as possible so as to avoid later memory loss. When done systematically and consistently, these documents can actually be subjected to formal desk review, or at the very least, summarized in a single country case study comparison matrix [IED #61].

INDEPENDENT EXPERT ASSESSMENTS

What they are and how OIOS-IED uses them
Independent expert assessments have been used by OIOS-IED to obtain expert review of products and activities of the UN Secretariat programme(s) being evaluated. This has included the convening of an expert panel to review flagship documents and the use of an individual expert to benchmark a key programme function against similar functions in national governments and the private sector. Owing to their expertise and their non-UN status, independent experts can carry considerable credibility as a source of external validation of (or counterpoint to) OIOS-IED's analysis.

As with any other type of systematic desk-based analysis, independent expert assessments must be carefully managed so as to ensure the integrity of the analysis. By virtue of their expertise, some reviewers might have prior contact with the programme in a way that biases their assessment negatively or positively. In other cases, experts might be harsher or more lenient in their appraisal than OIOS-IED would.

Because independent experts often come from different backgrounds and have different degrees of familiarity with evaluation methods, it is important to control for any potential bias. To do this, the evaluation team should develop a standard assessment tool that all reviewers use in their assessment [IED #62]. This ensures uniformity in what they are looking for. If used, such tools entail a responsibility on OIOS-IED's part to brief reviewers on the instrument and what is and is not expected of them. Wherever possible, OIOS-IED should also pilot the instrument being used in the review and control for bias by calculating an inter-rater reliability score to ensure all reviewers have a shared understanding of the criteria at play and how these are to be applied [📖 NCBI] [📖 MED-ED]. Any significant differences of approach can then be discussed and reconciled.

REMOTE MONITORING

What it is and how OIOS-IED uses it

While not typically used by OIOS-IED, another data collection method is remote monitoring. Remote monitoring is an approach used in contexts where there is restricted or limited access to the programme area. It has been employed by humanitarian agencies working in conflict areas, and can also be used in situations where limited resources prohibit travel. As traditional remote monitoring contexts often have missing or unreliable data, remote monitoring methods aim to both collect sufficient quantities of data, and to verify existing information from local partners, when in-person visits are not possible. While remote monitoring shifts risks and responsibility to local implementing partners, these organizations generally have a better understanding of the programme context and stronger ties to protective community structures. Local capacity building is often a key component of many remote monitoring plans.

Remote monitoring methods vary according to context, but might include: use of third parties for data collection and verification; beneficiary-led monitoring; cross-monitoring amongst local implementers; and correlation of data across sectors. In addition, widespread mobile phone usage has allowed for new and efficient methods of collecting data, including photo and video evidence. Technologies such as SMS and crowd-sourcing mapping platforms such as Ushahidi [📖 Ushahidi] can provide a large amount of information organized geographically or by need, and can be managed remotely.

While remote monitoring is commonly used in routine programme monitoring situations, similar methods apply for evaluation data collection as well, particularly when evaluations cover a UN Secretariat programme spanning a large number of countries. As external evaluators are less likely to have local partnerships, data collection methods relying on third-party monitoring and/or mobile technology might be more appropriate.

3.2   Choosing Data Collection Methods

Depending on the type of UN Secretariat programmes to be evaluated, OIOS-IED evaluation teams make use of any combination of different data collection methods.

Table 15 summarizes those data collection methods that are typically the most useful in evaluating each of the four types of UN Secretariat programmes' areas of work (Part I, Section 1.3).

Table 15: Most Useful Data Collection Methods, by Type of Secretariat Work

Data Collection Methods	Type of Work			
	Normative	Analytical	Operational	Internal Support Services
Interviews	Interviews with Member State representatives to determine UN Secretariat programme contributions to norm setting; Interviews with civil society representatives active in normative work; Interviews with programme staff to better understand their roles in norm-setting work.	Interviews with Member State representatives to determine UN Secretariat programme contributions to analytical work; Interviews with key users of the Organization's analytical work to determine its quality, relevance and usefulness; Interviews with civil society representatives active in analytical work; Interviews with programme staff to better understand their roles in analytical work.	Interviews with Member State representatives at country level to determine the programme's contribution to national priorities and international commitments; Interviews with key UN system partners at country level to determine the programme's contribution to UN operational activities; Interviews with civil society representatives active in similar operational work; Interviews with programme staff to better understand their roles in operational activities.	Interviews with members of the body being serviced (the clients) to determine timeliness and quality of service(s); Interviews with Member States of the inter-governmental body; Interviews with programme staff to better understand their roles in service.
Focus Groups	Discussions with programme staff to identify strengths and weaknesses in achieving targeted normative results; Discussions with partners to assess UN Secretariat programme contribution to norm setting.	Discussions with programme staff to identify strengths and weaknesses in achieving targeted results associated with analytical work; Discussions with partners to assess UN Secretariat programme contribution to results in shared analytical work.	Discussions with field staff to identify strengths and weaknesses in achieving targeted results, including Headquarters support to field operations; Discussions with civil society representatives and other partners to assess UN Secretariat programme contributions to targeted results in various areas of intervention; Discussions with members of beneficiary population(s) to exchange perspectives on what has gone well and what has gone less well.	Discussions with programme staff to identify strengths and weaknesses in achieving targeted support results; Discussions with internal UN clients to assess strengths and weaknesses in achieving targeted support results.

Surveys	Surveys of Member State representatives to determine UN Secretariat programme contributions to norm setting; Surveys of civil society representatives active in normative work; Surveys of programme staff to better understand their roles in norm-setting work; Staff time use and cost surveys of normative output production.	Surveys of Member State representatives to determine UN Secretariat programme contributions to analytical work; Surveys of key users of the Organization's analytical work to determine its quality, relevance and usefulness; Surveys of civil society representatives active in analytical work; Surveys of programme staff to better understand their roles in analytical work.	Surveys of Member State representatives at country level to determine the programme's contribution to national priorities and international commitments; Surveys of key UN system partners at country level to determine the programme's contribution to UN operational activities; Surveys of civil society representatives active in similar operational work; Surveys of programme staff to better understand their roles in operational activities. Field surveys of local population (beneficiaries and non-beneficiaries alike) to cull their knowledge, attitudes and behaviours related to the evaluand's work, and to assess coverage (including underserved subpopulations).	Surveys (client satisfaction surveys) of members of the body being serviced (the clients) to determine timeliness and quality of service(s); Surveys of Member States of the inter-governmental body; Surveys of programme staff to better understand their roles in service.
Direct Observation	Direct observation of inter-governmental meetings where norms are discussed and established.	Direct observation of international fora where analyses are discussed and reviewed.	Direct observation of operational activities.	Direct observation of meetings where service(s) are provided.
Desk Review	Collection and review of normative documents; Collection and review of national legislation relevant to normative to assess how norms have been incorporated into national and local government practices.	Collection and review of analytical documents;	Collection and review of documents, including planning documents, monitoring and self-evaluation reports.	Collection and review of documents prepared by the UN Secretariat programme providing the support services.
Independent Expert Assessments		Commissioning of independent expert assessments of the quality of the Organization's analytical work.		

3.3 Sampling

This section devotes attention to the important aspect of sampling that underpins the data collection effort and enables credible analysis.

Sampling is a method to obtain estimates on a large group, referred to as "the population" or "the universe", from a smaller sub-set of that group. Rather than collecting information from the entire universe (called "census"), sampling is more efficient - i.e., less costly and less time-consuming.

When using sampling to obtain information, OIOS-IED evaluation teams undertake the following basic steps:

- Define the universe - sampling begins by clearly setting the boundaries for the universe of people or items ("units") - i.e., defining clearly and precisely what is in and what is outside the universe;

- Construct the sampling frame - i.e., the complete set of units (people, documents, events, and so forth), each of which is assigned a unique number, and from which the sample is drawn. Ideally, the list will be identical to the universe. Should it depart significantly, the ability to generalize will be compromised;

- Select the sampling techniques, based on needs and capacity - i.e., the method of sample selection. When it is important to be able to generalize or make statements about the entire population, random sampling is the preferred method. Non-random sampling does not permit generalization to the universe, but might be less expensive to carry out and can be faster to complete;

- Calculate sample size for random samples - the appropriate sample size is a function of the population size as well as desired levels of confidence and precision; and

- Determine method for determining non-random samples.

There are two categories of sampling methods: random sampling and non-random sampling.

RANDOM SAMPLING

Random sampling is a technique that permits results to be generalized to the entire population, based upon the principle that each unit within the defined population has an equal chance of being selected, thus removing selection bias. Each unit in the population is assigned a unique identification number. Units are then selected using a random number generator [📖 Stat Trek].

OIOS-IED commonly uses one of the following three types of random samples:

- Simple random samples
- Stratified random samples
- Random cluster samples

Simple random sampling

Simple random sampling [📖 BetterEvaluation] is easiest to use. It is a straightforward sampling type involving a random selection of a pre-determined number of units from a sampling frame. It is used when the primary objective it to make inferences about the whole - homogenous - population rather than about specific subpopulations.

Stratified random sampling

Frequently, it is important to include in a sample specific groups that might otherwise be poorly represented. To conduct stratified random sampling [📖 BetterEvaluation], the sampling frame is divided into distinct groups based on meaningful characteristics, or "strata", and a random sample is selected from each group (or "stratum"). This ensures enough sampling units to be able to generalize to each stratum, as well as to the total

population. In stratified random samples, it is necessary to weight the data during the analysis in order to compensate for the stratification- not all groups might have the same size. If, for example, one group is twice as big as another, the results from that group should have double weight.

Random cluster sampling

Another variation of the simple random sample is the random cluster sample drawn from naturally occurring clusters of the unit of analysis, and often used in the absence of a complete list of everyone/everything in the universe. For example, an evaluator might want to sample staff in a particular subprogramme, and might not have a complete list of every staff member in the subprogramme, but instead has a list of every division. It is therefore possible to select divisions randomly and then randomly select staff members within each of the selected divisions. The divisions represent a cluster of staff members.

Sampling with or without replacement

Another choice available in sampling is the option of sampling with or without replacement. When sampling with replacement, a unit can be selected and used more than once. In this case, the population from which the sample is drawn is regarded as infinite. When sampling without replacement, a unit is selected and used only once. Sampling without replacement is the most commonly used method of sampling within OIOS-IED.

Cohort replenishment

Replenishing samples, also referred to as cohort replenishment, is a method used in longitudinal or panel surveys where a group of units are tracked and measured over time. As time elapses, samples of this type are subject to attrition or loss of units from the initial sample. This tends to be cumulative over time and, if sizeable, can compromise the sample's ability to represent the initial universe. It can also be a source of statistical error as the number diminishes and its power to infer is reduced. A method

for solving this problem is to replenish the units lost to attrition in successive waves. Any given replenishment sample will be representative of the population at the time of data collection for the new wave to which they correspond, and not the original population.

CALCULATING RANDOM SAMPLES

Random samples contain errors due to the basic condition that not all units in the universe are being used. Two concepts define this error - i.e., confidence and precision, both of which determine the size of the sample used.

Confidence level

The first question to address is how confident the evaluator wants to be that sample results are an accurate estimate of the entire population. The greater the desired level of confidence, the larger the sample size needed. The OIOS-IED minimum confidence level for surveys is 90 percent - the evaluator should be at least 90 percent certain that sample results are an accurate estimate of the whole population.

Sampling error

The second question is how precise the estimates should be. This plus-or-minus figure is sometimes called confidence interval, margin of error or level of precision, and is often seen when results from polls are reported. For example, a poll might reveal that 48 percent of individuals favour raising taxes and 52 percent oppose raising taxes, with a margin of error of +/- 3 percentage points. What this means is that if everyone in the population were asked, the actual proportions would be somewhere between 45 percent to 51 percent in favour of raising taxes and 49 percent to 55 percent opposing a tax rise. The greater the desired level of precision, the larger the sample size needed. OIOS-IED typically accepts a sampling error of up to +/- 5 percentage points.

Sample size

Determining the appropriate sample size depends on population size as well as the required levels of confidence and precision. To ensure data reliability, OIOS-IED evaluators make use of online tools [📖 Creative Research Systems] to calculate sample size and margin of error.

There are a number of external resources that can prove useful in undertaking random sampling [➔ World Bank. 2009a] [➔ Gonick et al. 1993a] [➔ Urdan. 2010a] [📖 BetterEvaluation] [📖 Sophia] [📖 Stattrek].

NON-RANDOM SAMPLING

A common misconception is that systematic sampling is the sole domain of quantitative analyses. While OIOS-IED most commonly makes use of sampling in connection with surveys and interviews, its evaluation teams also sample documents for desk review, countries for inclusion as case studies, and project sites to visit within these countries.

It might not always be possible to use a random sample for reasons such as lack of resources, lack of time or the absence of a convenient sampling frame. In such situations, OIOS-IED evaluators turn to non-random sampling. Anytime the evaluation seeks to select cases for analysis (whether individuals, documents, countries or sites), in which it is not necessary, possible or desirable to randomly select cases for analysis, this is called a non-random sampling method.

There are two main types of non-random sampling methods:

- Convenience sampling [📖 BetterEvaluation] - selection of cases because they happen to be available. A good example of convenience sampling is the intercept interview, whereby individuals are selected for interview as they enter or exit a programme delivery site - e.g., clinic, food distribution site;

- Purposive sampling [📖 BetterEvaluation] - deliberate selection of cases because the evaluation team desires specific information that they know the case can offer. Cluster analysis used in country case study selection is one such example. "Snowball" samples are another, which are often used in developing sample lists of substantive experts - an initial group of such experts is approached, included in the sample, and asked for further names for inclusion in the sample; these further contacts are then targeted and in turn asked for further names, and so on.

In this type of non-random sampling, OIOS-IED evaluators select a small set of units from the population in order to make statements about that particular sub-set. Although ideally there are no obvious differences in the characteristics of the sample and the population, such as age or gender composition, evaluators should not generalize the results of the information obtained by non-random sampling to the general population, in the same way as using a randomly-based sample, as non-random samples can be subject to all types of bias.

When using a non-random sample, results should be reported in terms of the sample of respondents, rather than the corresponding universe. For example, "70 percent of the mission staff in the interview sample reported satisfaction with living conditions." This differs from stating "70 percent of mission staff reported satisfaction with living conditions."

Because the circumstances prompting the use of non-random sampling are varied, OIOS-IED teams are advised to consult the many external resources on this subject. [➔ Henry. 1990a] [➔ Daniel. 2011a] [➔ Levy et al. 2013a] [📖 Socialresearch] [📖 ABS]

Step 4: Data Analysis

| Step 1: Announcement | Step 2: Evaluation Design | Step 3: Data Collection | Step 4: Data Analysis | Step 5: Report Preparation | Step 6: Dissemination | Step 7: Housekeeping | Step 8: Tracking Follow-Up |

Data analysis is the stage at which the OIOS-IED evaluation team processes the data collected in order to answer the evaluation questions (Part II, Section 2.5). It is critical to conduct high-quality and rigorous data analysis if evaluation findings, conclusions and recommendations are to be soundly based on evidence and persuasively argued.

In reality, data collection (Part II, Step 3) and data analysis frequently overlap to some extent. For instance, while data are still being collected through telephone interviews, the OIOS-IED evaluation team might begin to analyze survey responses and written documents. Similarly, the results of individual analysis, such as a specific desk review, may be discussed in "real time" during the data collection stage itself. Nonetheless, the data analysis stage constitutes the formal step in the evaluation process in which all of the data are brought together and viewed holistically against the evaluation questions.This sub-chapter has seven main sections:

- OIOS-IED Data Analysis Process (Part II, Section 4.1)
- Assessing Data Validity and Reliability (Part II, Section 4.2)
- Data Compilation (Part II, Section 4.3)
- Analysis of Different Data Sets (Part II, Section 4.4)
- Overall Evaluation Data Analysis and Triangulation (Part II, Section 4.5)
- Analysing Qualitative Data (Part II, Section 4.6)
- Analysing Quantitative Data (Part II, Section 4.7)

4.1 OIOS-IED Data Analysis Process

Before conducting any data analysis, OIOS-IED evaluation teams discuss how to approach the analysis of different data sets. They develop an overall evaluation data analysis plan rooted directly in the evaluation design matrix (Part II, Section 2.5) and used to systematically and strategically guide the process by which data are analyzed. This helps to minimize inter-evaluator variability that might otherwise reduce the reliability of the process.

The OIOS-IED Quality Assurance System (QAS) (Part I, Chapter 3) includes a checklist for data analysis [IED #24]. This checklist helps OIOS-IED evaluation teams ensure that critical items described in this section are included in the data analysis plan.

4.2 Assessing Data Validity and Reliability

To maximize the credibility of their analysis, OIOS-IED evaluation teams double-check that data are valid and sufficiently reliable before they analyze each primary and secondary evaluation data set.

One visual analogy often used to differentiate validity and reliability is that of a dartboard. Figure 10 serves to visually illustrate this analogy.

Figure 10: Validity and Reliability Dartboard
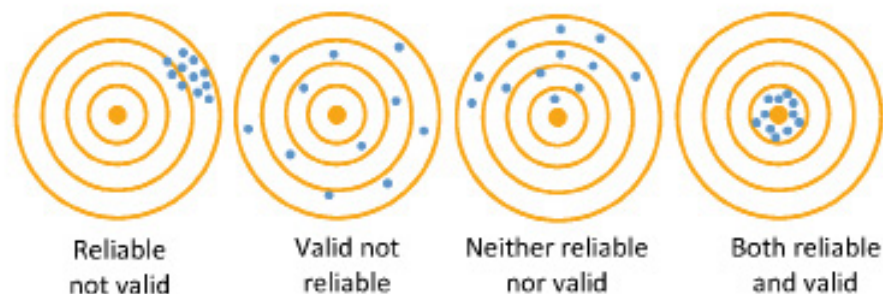
In Figure 10, the "bull's eye" represents the concept the evaluation team is interested in measuring and the dart marks represent the data points in a given data collection. The farther "off" from perfect measurement , the farther one's dart is from the centre. In the leftmost dartboard in Figure 10, the evaluation team is hitting the target consistently, but missing the bull's eye – i.e., they are consistently and systematically measuring the wrong value for all respondents. This measure is reliable, but not valid. The next dartboard shows dart marks that are randomly spread over the board. The evaluation team seldom hits the centre of the target, but on average is getting the right answer for the group as a whole. In this case, the evaluation team is getting a valid group estimate, but is inconsistent. The third dartboard shows a case where the dart marks are spread across the target and are consistently missing the bull's eye. The evaluation team's measure is neither reliable nor valid in this case. Finally, the rightmost dartboard depicts the ideal scenario - the evaluation team is consistently hitting the centre of the target with measurements that are both reliable and valid.

DATA VALIDITY

Data are considered valid when they truly represent what is being measured. There are two main forms of validity - internal and external validity. Both are equally important. Although drawn from the literature on experimental designs (Part II, Section 2.7), these concepts can help OIOS-IED evaluation teams approach their data analysis in a more robust way. Anything that compromises either of these forms of validity is referred to as a threat to validity.

- Internal validity - threats to internal validity compromise our confidence in saying that a relationship exists between the independent and dependent variables (e.g., by suggesting that there is insufficient evidence for causal claims being made, owing to alternative explanations to explain the dependent variable); and

- External validity - threats to external validity compromise our confidence in stating whether the study's results are applicable to other groups (e.g., by calling into question the sampling methods, response rates or other aspects of the data collection effort).

There are numerous external resources that speak to validity and specific tools for countering threats to validity [➔ Cook et al. 1979a] [➔ Gonick at al. 1993a] [📖 Socialresearch].

For OIOS-IED's purposes, all data collected should be examined for validity at the beginning of the data analysis stage. Key questions for the evaluation team to ask include the following:

- When using random samples, were response rates and sample sizes sufficiently high to warrant a 95 percent level of confidence in making statistical claims?

- When using non-random samples, are results to be reported in terms of the respondents and not generalized to the population?

- Are secondary data collected for the analysis sufficiently valid? What, if any, limitations do they entail for the team's analysis?

- When considering evidence to be used for causal arguments, such as arguments related to the programme's effectiveness or ineffectiveness in contributing to results, what, if any, alternative explanations might be at play other than the "causal" variable being explored?

DATA RELIABILITY

Data - both qualitative and quantitative (Box 7) - should be examined for reliability at the outset of the data analysis stage. Data are considered reliable when they are accurate and complete. Data accuracy and completeness are achieved only when the data collection tools used to generate these data, and the process used to collect them, produce stable and consistent results across all evaluators and all stakeholders.

There are numerous external resources that speak to reliability [➜ Cook et al. 1979a] [➜ Gonick at al. 1993a] [📖 Socialresearch].

Each of these aspects of data reliability - i.e., accuracy and completeness - is discussed in turn here.

Data accuracy

Data accuracy is the extent to which data were obtained in a clear and well-defined manner from different datasets across cases, space and time as well as the extent to which they truthfully reflect facts and figures such as dates, percentages and numbers of persons interviewed. Testing for accuracy helps establish that any variations in the data originate from differences in the actual situation. Lines of exploration include:

- Did evaluation team members maintain an independent, balanced and objective attitude and approach toward each and every interviewee?

- Did survey respondents understand terms or questions in the same way or is it possible that they made different interpretations?

- Were observational guides and desk review instruments used consistently by all evaluation team members, ensuring maximum fidelity to the instrument?

- Did all enumerators use the field-based survey in the same way, ensuring maximum fidelity to the survey instrument?

- Were coding rules consistently applied?

- If appropriate, were confidence intervals for key survey estimates analyzed?

- Wherever appropriate, were appropriate weighting schemes applied to adjust for over- or under-sampling?

Data completeness

Sometimes attempts to gather data are unsuccessful. This could be because data were found to be unavailable or unobtainable. At other times, individuals might be unavailable or unwilling to participate in surveys or interviews. Such gaps in the data are relevant if they prevent the planned quality, quantity or representativeness of the data from being achieved. It might mean that results could be biased towards the views of those who participated, which might not be the same as the views of those who did not participate.

One particular problem associated with incomplete data is that of non-response. As individuals who have not responded might be distinctly different from those who have, non-response error is an issue that OIOS-IED evaluation teams should address when response rates are low.

ADDRESSING RELIABILITY CONCERNS

A number of steps can be undertaken at the data analysis stage to strengthen data re-liability. For example, once the extent of any gaps or inadequacies in data are known, it might still be possible to plan and undertake some supplementary data collection depending on budget and other resources.

In cases of non-response, the evaluation should consider undertaking a non-response analysis to test for the likelihood of bias in the sample. Non-response analysis is a comparison of respondents' demographic attributes against those of non-respondents, to the extent any of these is known. If there is little or no evidence that these groups differ, then one can more safely assume that the distribution of results is likely to be substantially unchanged if the full results are known.

Another, more intricate means of adjusting for non-response (and not usually used in OIOS-IED) is that of data imputation. In data imputation, the evaluation team assigns values to missing cases using any number of techniques, then controlling for the imputed values as a dummy variable. Given its level of so-phistication, imputation might require consultation with an expert statistician.

For systematic analyses undertaken by the evaluation team, such as desk reviews and direct observations, an inter-rater reliability score can be calculat-ed to ensure that all evaluation team members undertaking the analysis were applying the data collection instruments consistently [📖 NCBI] [📖 MED-ED].

As well as being considered during the data analysis itself, any substantial inad-equacies encountered should also be declared in the methodology limitations section of the evaluation report (Part II, Step 2).

Box 7: Data Cleaning - Preparing Quantitative Databases for Analysis

Data cleaning [📖 BetterEvaluation] is an important analysis preparatory step in all data collection instruments that include quantitative elements. Put simply, data cleaning is a means of boosting the accuracy of the dataset by removing any suspicious data and handling any missing data. At minimum, data cleaning includes the following steps:

- Analysing and adjusting for non-responses;
- Removing duplicate cases (e.g., multiple survey submissions from the same IP ad-dress);
- Examining extreme outlier cases to determine the extent to which they might skew results, or even represent bogus responses;
- Assessing attrition rates at different points in a survey, and determining whether to throw away or use these partially completed surveys; and
- Weighting the dataset to adjust for any over- or under-sampling, such as when stratified random sampling is used.

4.3   Data Compilation

Data compilation involves gathering the complete data into files on the OIOS-IED shared drive - taking into account OIOS-IED's confidentiality policy - e.g., by password-pro-tecting files indicating individual stakeholders' identities - in order to ultimately make available a comprehensive and quality-checked evidence base for the evaluation. The transparent availability of such a record is critical should others wish to verify or examine the evidence compiled in support of the evaluation results.

Data compilation goes beyond mere file organization, however. It is also about "aggregating up" the raw data into more manageable sources of information that are more readily analysable by the evaluation team. This can already begin during data collection. For example, team members' individual aide mémoires,

document summaries and other items of relevance to country case studies can be summarized into a single spreadsheet or other file for ease of comparison by the full team during data analysis. Frequency distributions for all survey data, desk reviews, direct observations and other methods can be distilled down into summary tables once they are completed, so that they are ready for data analysis. In addition, important individual data points, observations or questions of relevance to the analysis can be stored in a "parking lot" file, to be used later on during data analysis.

4.4   Analysis of Different Data Sets

Once individual data sets have been quality-assessed and compiled, analysis helps to make sense of what the different strands of evidence are saying. Analysis is conducted against each of the evaluation questions. It is also useful to add an "other" category to enable unexpected issues to become visible.

As illustrated in Figure 11, this step typically results in insights and preliminary results based on:

- For structured and semi-structured interviews - coded responses to individual questions;

- For surveys - tabulation of closed-ended responses, including possible cross-tabulations between two variables, and coding of open-ended responses;

- For focus groups - coding of areas of stakeholder convergence and divergence;

- For structured and semi-structured direct observation - tabulation of closed-ended data and coding of open-ended results arising from use of the observational instrument; and

- For desk review - tabulation of closed-ended data and coding of open-ended results arising from use of the desk review tool.

Graphics can also be used to supplement and depict the written analysis. Bar charts show differences across categories or across time. Pie charts show the distribution of a variable (expenditure, staff, and so forth) among different categories. Maps show the distribution of a variable over a geographic area.

The options for rendering data visually are virtually endless, and there are numerous external resources [→ Tufte. 1990a] [→ Tufte. 1997a] [→ Tufte. 2001a] [→ Tufte. 2006a] to help OIOS-IED evaluation teams think creatively about how to do so.

Compelling data visualization can not only help the evaluation team make better sense of its various data streams more quickly. Later, during the report preparation stage (Part II, Step 5), it can also help convey the evaluation story in an engaging way, thus paving the way for higher utilization of the evaluation report.

All outputs arising from these data analyses - whether tabulations, coding results or graphics - are placed on the OIOS-IED shared drive, with any identifiers removed that might connect the names of individual stakeholders to the data. In addition, it is often useful to produce a portfolio of evidence for each team member's review - i.e., a binder including all summaries of the raw data and, where applicable, the raw data themselves.

4.5   Overall Evaluation Data Analysis and Triangulation

The moment at which the OIOS-IED evaluation team comes together to begin reviewing the various data sources against the evaluation questions is the most crucial part of the data analysis. In OIOS-IED this typically involves one or more brainstorming sessions, in which the team members, having read the portfolio of evidence separately on their own, meet to review and synthesize the collective evidence together, and to uncover what the evidence is telling them with regard to the key evaluation questions.

Figure 11: The Data Analysis Process
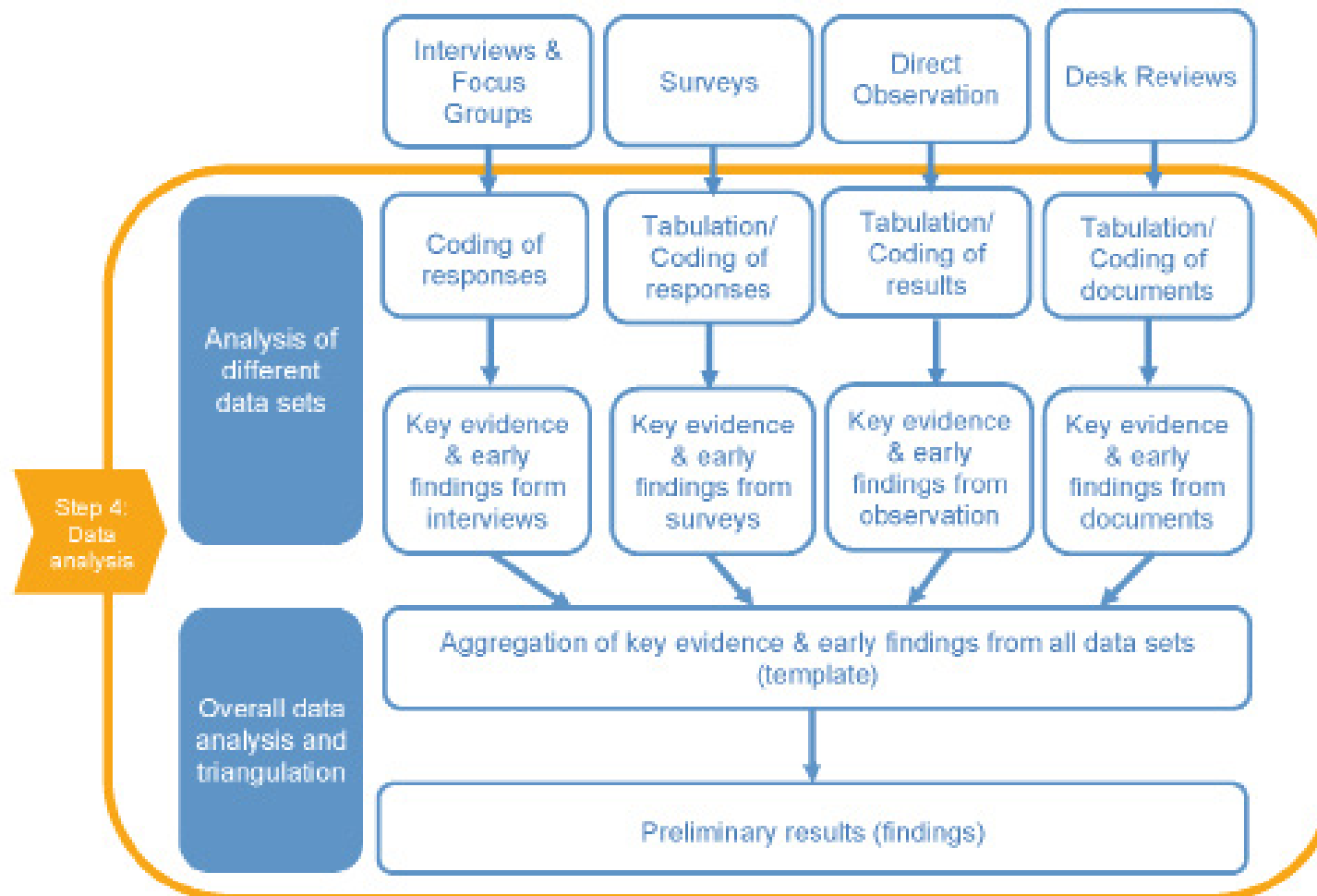
At the end of these sessions, the team converges around the main "story" the evidence is telling about the evaluand's relevance, effectiveness, impact and efficiency.

In practice, an integrated analysis begins with the creation of an evidence summary table [IED #25] with the key evaluation questions and any subsidiary questions. Against these are listed what the different data sets say (interim analysis). This enables the detection of patterns across different data sets, which is called triangulation [📖 BetterEvaluation]. Overall evaluation data analysis can also pinpoint differences and contradictions among data sets that require reconciliation and/or further exploration.

4.6   Analysing Qualitative Data

Analysis of qualitative data [➔ World Bank. 2009a] [📖 BetterEvaluation] is called content analysis. Content analysis is applied in order to review different types of documentation, both primary and secondary information sources, in a systematic way, in order to bring order out of the wide array of qualitative data collected in the evaluation. The evaluator might be seeking to identify patterns in the data, including the frequency of concepts or the relationships among them, as represented by words or phrases mentioned in interviews or documents. Ultimately this process of pattern-seeking helps the OIOS-IED evaluation team draw inferences.

OIOS-IED uses two standard types of content analysis, i.e. conceptual content analysis, which is concerned with the frequency of concepts, and relational content analysis, which focuses on relationships among concepts.

CONCEPTUAL CONTENT ANALYSIS

Conceptual concept analysis involves these steps for each evaluation question:

- Decide on the level of analysis - predetermine which concepts to search for; they must be mutually exclusive. Build in flexibility to accommodate other concepts that might emerge reflecting important concerns;

- Establish rules for detecting concepts - often, it might not be clear which phrases or words represent certain concepts. Moreover, there might be implicit references;

- Develop a set of coding rules - explicit rules for coding are critical to ensure the quality of information. Coders should be trained in the consistent application of these rules because of implications for data reliability;

- Conduct text coding - coding is done by one or more coders, either manually or with computer assistance. Dedicated software packages can greatly speed up the process; and

- Conduct analysis - tabulate frequencies for each concept and reveal patterns and commonalities.

Focusing on how often a concept occurs (frequency) limits the level of analysis to the quantitative nature of these concepts. Relational concept analysis also begins by predetermining concepts, but attempts to attain a higher level of interpretation by establishing explicit and/or implicit relationships between concepts.

RELATIONAL CONCEPT ANALYSIS

Conducting a relational analysis consists of the following steps:

- Identify the evaluation question to be answered;
- Conduct text coding;
- Explore the relationships between concepts for characteristics such as frequency, direction and strength of association;
- Code different types of relationships;
- If appropriate, apply statistical analysis; and
- Map out relationships.

There are a number of software packages that can help with analysing qualitative data. Word processing packages such as Word, or spreadsheets such as Excel, have features that can assist with searches, indexing and manipulating the material for analysis. Additionally, there are software packages specifically designed for handling these types of data, generally referred to as Qualitative Data Analysis Software, or QDAS, and include NVivo, which OIOS-IED currently uses. These packages assist with coding a vast amount of material using a number of different strategies.

Although an extremely useful source of rich qualitative data - and of quantitative data that summarizes patterns in the qualitative data - content analysis is challenging to do well, in that it:

- Can be very time-consuming;

- Depends on timely availability of documents, good notes and quality of content;

- Entails a risk of evaluator bias, even when document quality is high, in terms of the subjectivity that affects coding when coding rules are inconsistently applied; and

- Requires highly in-depth analysis in cases where concepts and codification might not be simple or straightforward to identify in the text (e.g., when documents contain implicit rather than explicit references to what the evaluation team is looking for).

Inconsistent application of coding rules or pre-conceived positions can easily affect the coding and interpretation of data, often more easily than with quantitative data. OIOS-IED evaluation team members work together throughout the entire process to help prevent a biased interpretation at the end. They develop a system of validation for each stage of the analysis, including the end point when results are summarized to provide answers to evaluation questions. While this is not a guarantee, it can be an effective way of minimizing introduction of bias into evaluation results.

4.7   Analysing Quantitative Data

Quantitative data analysis [➔ World Bank. 2009a] [📖 BetterEvaluation] summarizes numerical information collected during the course of an evaluation. In its simplest form, quantitative data analysis merely involves tabulating the numbers or percentages on phenomena of relevance to the evaluation questions. Such quantitative data can be useful to the analysis regardless of the type of sample drawn. For example, counting the number of resource mobilization policies a UN Secretariat programme has had in place over a strategic planning period, the percentage of all staff that have been trained in gender mainstreaming, or calculating the time to deployment of a support unit to the field in a given year, are all potentially valuable sources of quantitative information.

When random samples (Part II, Section 3.3) are drawn to make generalizations about a universe, OIOS-IED's quantitative data analysis might also involve the application of statistical methods or tests to numerical data as well. These can range in complexity from very simple descriptive methods to very complex multivariate analyses. Statistical concepts applied to OIOS-IED's work fall into three general categories:

- Descriptive statistics - show the current situation or condition;
- Associational statistics - look at how variables change together; and
- Deterministic (or causal) statistics - look at how the change in one variable affects another.

There are a number of established statistical software packages that can perform a multitude of tests for an evaluation. Given the level of sophistication, associational and deterministic statistics might require consultation with an expert statistician.

DESCRIPTIVE STATISTICS

Descriptive statistics [📖 BetterEvaluation] are the most commonly used method for quantitative data analysis, and the most commonly used in OIOS-IED, with the help of Excel or SPSS Statistics. Descriptive statistics are statistics used to describe and summarize quantitative data, mostly in terms of frequencies or percentage distributions - e.g., 33 percent of the respondents are male and 67 percent are female.

Typically, data are summarized using three different kinds of descriptive statistics:

- Measures of frequency - summarize the numbers or percentages of cases exhibiting a certain attribute, responding in a certain way, and so on;

- Measures of central tendency - summarize the "typical" or "average" response across all of the cases included in the analysis; and

- Measures of dispersion - look at how close or far away cases tend to be from the "typical" or "average" of the group as a whole.

Measures of frequency

Frequency is the number of times a data value occurs. For example, if 25 households in a community have two children, the score of two has frequency of 25 (Table 16). A frequency distribution table [📖 BetterEvaluation] makes it easier to understand a data set and allows for its graphical representation.

Table 16: Example of Frequency Table (number of children per household)

Number of children per household (x)	0	1	2	3	4	5	x > 5
Frequency (number of households with x children)	7	18	25	19	13	8	4

Measures of central tendency

There are three basic measures of central tendency [📖 BetterEvaluation], commonly referred to as the 3 Ms - the mean, the median and the mode. The mode represents the most frequent response or characteristic. The median represents the mid-point or middle value in a distribution. The mean represents the arithmetic average. The two most commonly used statistics are the mean and the median.

It is important to understand that the preferred use of these three measures of central tendency depends upon the type of data available:

- Nominal data - i.e., names or categories, such as gender, religion, country of origin) -> mode;

- Ordinal data - i.e., data that can be placed on a scale that has an order to it; e.g., scales that go from "most important" to "least important," or "strongly agree" to" strongly disagree" -> mode or median; and

- Interval/ratio data - i.e., real numbers with a zero point and fixed intervals that can be divided and compared to other ratio numbers -> mode, median or mean (Table 17).

Table 17: Example of Mode, Median and Mean

	Data (interval/ratio)	Data Analysis
Mode	9, 15, 15, 15, 17, 23, 27, 40, 44, 44	The mode is 15. (15 occurs more often - at 3 times – than any other value)
Median		The median is 20. (Add the 2 middle numbers and divide by 2 -> 17 + 23 = 40 ÷ 2 = 20
Mean		The mean (average) is 24.9. (The sum of 249 divided by 10)

There is one further qualification relating to use of the mean - for interval/ratio data, the choice will also depend on the distribution. If it is a normal distribution, the mean, median and mode should be very close. In this case, the mean would be the best measure of central tendency. However, with few very high scores or few very low scores, the mean will no longer be close to the centre. In this situation, the median is a better measure of central tendency.

Measures of dispersion

The most commonly used measure of dispersion [📖 BetterEvaluation] for interval or ratio data is the standard deviation. Standard deviation is a measure of the spread of the scores on either side of the mean.

The more the scores differ from the mean, the larger the standard deviation will be - if everyone scores 75 on a test, the standard deviation would be 0. If everyone scores between 70 and 80 (mean 75), the standard deviation would be smaller than if everyone scored between 40 and 90 (mean 75). In summary:

- Small standard deviation = not much dispersion
- Large standard deviation = lots of dispersion

ASSOCIATIONAL STATISTICS

The most common associational statistic is the correlation coefficient. Correlation simply measures whether two variables "move together" - i.e., whether when one increases, the other tends to increase, and vice versa. Correlation is not causation, however. It does not indicate whether one of the variables causes the other, and if so which, or whether both are caused by a third variable.

Examples of correlational analysis that might be useful in OIOS-IED evaluations include the following:

- The correlation between a country office's overall financial size and the level of financial investment in evaluations of its work;
- The correlation between the source of a global programme and the speed at which the programme's programme review committee approved it;
- The correlation between respondents' staff level and their support for a change management initiative, as self-reported in a staff survey; and
- The correlation between a beneficiary's gender and the speed at which s/he is served in a project implementation site, as recorded in OIOS-IED's direct observations.

Contrary to what is commonly thought, correlation analysis is not as straightforward as it seems. The specific correlation coefficient to be calculated depends on the specific variable type of each of the two variables being examined. OIOS-IED has developed a brief fact sheet providing guidance on which correlation statistic to use for each set of variables being examined [IED #63].

DETERMINISTIC (OR CAUSAL) STATISTICS

As noted above, association does not mean causality. To determine causality, another type of statistics - deterministic statistics - is needed. Given the complex nature of the UN Secretariat progammes OIOS-IED evaluates, as well as the limited availability of existing quantitative data on the programme to analyze, OIOS-IED rarely uses deterministic statistics to establish statistically valid causation between the evaluand and its work (as an independent variable) and its targeted objectives (as the dependent variable). Instead, it is much more common for OIOS-IED to discuss "contribution" rather than "attribution" when exploring the evaluand's role in influencing change.

Deterministic analytical methods attempt to establish a causal relationship between two or more variables. The most common deterministic method is simple regression analysis, which explores the causal relationship between one or more independent variables (i.e., variables thought to contribute to causing something else) and a dependent variable (i.e., the variable whose causes one wishes to explore). When only one independent variable is thought to be influencing the dependent variable, the form of regression is called bivariate regression. Since real-world phenomena are almost always more complex than any single variable can explain, however, the most common form of regression analysis is multiple-regression analysis. Multiple-regression analysis (also known as multivariate analysis) explores the causal relationship between two or more independent variables and the dependent variable of interest.

The standard form of multiple regression analysis is Ordinary Least-Squares (OLS) regression, which assumes that both the independent and dependent variable are continuous, as well as a number of other assumptions. There are numerous variations of this standard OLS form of regression, however, which address violations to these various assumptions. (One of these - i.e., log-linear regression - allows the use of ordinal and categorical dependent variables.)

STATISTICAL INFERENCE

An important issue in quantitative data analysis is that of statistical inference. Statistical inference can become an issue when data from a random sample are used. The main question is whether results can be generalized to the population based upon the sample data. The issue of statistical inference is relevant to descriptive, associational and deterministic methods, whenever random sample data are involved.

Statistical inference is used to make an estimate about a population based on a random sample selected from that population. Whenever sample data are used, a major concern is whether the results are a function of the sample itself rather than a true picture of the population. If a different sample had been selected, would the results be similar? Statisticians have developed tests to estimate this, called statistical significance tests [➔ Gonick at al. 1993a] [➔ Chow. 1996a] [➔ Kline. 2004a] [📖 WikiHow] [📖 Researchrundowns]. They estimate how likely it is that results obtained in the analysis of the sample data are valid or whether they were obtained by chance alone.

All tests of statistical significance are partly based on sample size. They also assume that this sample size is rooted in a sufficiently high response rate. If the sample is very large, even small differences are likely to be statistically significant. Two of the more common tests are the chi-square and the t-test. The chi-square is one of the simplest statistics. It tests whether there is some underlying variance between two groups on a given attribute (e.g., the difference between women's and men's left-versus-right-handedness) that should be explored further in more sophisticated statistical tests. The t-test tests whether the differences between two groups on a given dependent variable, such as females' versus males' or management's versus staff's self-reported opinion that their programme is heading in the right direction, is statistically significant.

In any statistical test, evaluators typically set the benchmark for statistical significance at the 0.05 level ("p-value"), establishing that there is at least 95 percent certainty that the sample results are not due to chance. Given the importance of sample size and response rate in statistical inference, it is important that OIOS-IED evaluation teams, in their data analysis and report preparation (Part II, Step 2), not make statistical inferences based on insufficient levels of either of these. Small samples and low response rates should be indicated in the methodology section of the evaluation report.

In addition, just because there is a statistically significant difference does not automatically mean that the difference is important. The importance of analytical results is ultimately a judgment call.

Step 5:   Report Preparation

| Step 1: Announcement | Step 2: Evaluation Design | Step 3: Data Collection | Step 4: Data Analysis | Step 5: Report Preparation | Step 6: Dissemination | Step 7: Housekeeping | Step 8: Tracking Follow-Up |

Report preparation is the culmination of the data collection (Part II, Step 3) and data analysis (Part II, Step 4) undertaken for the evaluation. It is the stage at which the OIOS-IED evaluation team transforms its analysis into a clear, coherent and compelling story about the UN Secretariat programme's relevance, effectiveness, impact and efficiency. In order for OIOS-IED evaluation reports to achieve the Division's end objective emphasized in its PIP (Part I, Section 1.2) - i.e., the production of timely, objective, credible and relevant information that its stakeholders use to improve programme performance - it is essential that OIOS-IED's reports are consistently timely and of the highest quality.

The OIOS-IED Quality Assurance System (QAS) (Part I, Chapter 3) includes a checklist that helps evaluation teams keep track of the important elements of the report drafting process [IED #26]. In addition, OIOS-IED's Report Tracking Tool is an essential platform for ensuring that teams meet all major milestones of the report drafting stage [IED #64].

This sub-chapter has two main sections:

- OIOS-IED's Report Drafting Process (Part II, Section 5.1)
- Writing High-Quality Evaluation Reports (Part II, Section 5.2)

5.1   OIOS-IED's Report Drafting Process

The report drafting process commences even before the evaluation team begins drafting. The series of brainstorming sessions it undertakes during the overall evaluation data analysis stage (Part II, Section 4.5) should culminate in one or more structured sessions, with the Section Chief's full participation, to arrive at:

- An evaluation report outline, with agreement on type of evidence to be used and rough word limits for each section;
- Draft result statements - i.e., preliminary results; and
- A plan for drafting the report, including team member assignments and internal drafting, sharing and reviewing deadlines.

While all members of the OIOS-IED evaluation team are given drafting assignments, it is the responsibility of the Team Leader to consolidate individual sections into a single cohesive and logical report. The Section Chief may also draft sections of the report. Additionally, s/he provides overall supervision, feedback and guidance for drafting the report and must clear the final draft before it goes for review to OIOS-IED's Director and Deputy Director.

Prior to this, however, the team must conduct informal briefings (in person, by video-conference or by teleconference) with the OIOS-IED Director and Deputy Director, the evaluand and the OIOS USG. Subject-matter experts and advisory panel members (Part II, Section 2.3) are also consulted during the drafting process.

The briefing, review and clearance steps for OIOS-IED draft evaluation reports are summarized in Box 8.

Box 8: Briefing, Review and Clearance Steps for OIOS-IED Draft Evaluation Reports

1. Evaluation team briefs OIOS-IED's Director and Deputy Director on preliminary results;

2. Evaluation team briefs evaluand on preliminary results;

3. Evaluation team briefs OIOS USG on preliminary results;

4. Section Chief reviews and clears draft report for review by OIOS-IED Directorate (this typically takes several draft versions);

5. OIOS-IED's Director and Deputy Director review and clear draft report; and

6. OIOS USG reviews and clears draft report.

During steps 1 to 5 above, an informal version of the draft evaluation report is shared with the evaluand, subject-matter experts and advisory panel(s) for informal feedback. Once the draft report has been reviewed and approved by the OIOS USG (step 6) - and final changes made based on his/her comments - the final draft report is shared for formal comments with the Head of Department of the entity evaluated/inspected (step 7) [sample memo IED #65 - 66]. Once comments are received, they are fairly considered and incorporated where appropriate (step 8). The full text of the memorandum containing the formal comments is appended to GA reports.

After final review and approval by the Section Chief and OIOS-IED's Director and Deputy Director (step 9), the report is finalized (if it is a non-GA report) and disseminated (Part II, Step 6) or sent to the Office of the OIOS USG for clearance (if it is a GA report).

The latter is subsequently sent to the DGACM by the official "slot date" for formal editing and translation [form, checklist, memo IED #67 - 69]. The slot date is at the end of February for reports submitted to the GA Fifth Committee and at the end of March for reports submitted to the CPC. Failure to meet the slot date results in an official designation of a late report. Both non-GA and GA reports require the evaluand to

submit an action plan for implementing recommendations that includes the specific action(s) to be taken as well as the target date(s) for completion [IED #70 - 71]. Figure 12 illustrates the sequence of steps for vetting GA reports.

5.2   Writing High-Quality Evaluation Reports

OIOS-IED evaluation reports are geared towards busy readers. However complex the issues addressed, they should be easy to read and understand for those readers who are not experts on the issues being discussed. OIOS-IED reports should be written clearly, concisely and convincingly. Abbreviations and acronyms should be avoided to the extent possible. Sentences must be precise and neutral.

There are a number of external resources that summarize good practice for writing evaluation reports [➔ World Bank. 2009a] [➔ UNEG. 2010b] [📖 BetterEvaluation]. The OIOS-IED Quality Assurance System (QAS) (Part I, Chapter 3) includes a checklist that also addresses the contents of evaluation reports, including recommendations [IED #26].

Non-GA reports do not have formal word restrictions. However, it is considered good practice to limit their length to no more than 8,500 words, since shorter reports are more likely to be read, more accessible and more compelling. Inspections or evaluations with a specific GA mandate must follow a prescribed format. GA reports are limited to 8,500 words, including footnotes and annexes (the annex containing the evaluand's formal comments is exempted from inclusion in the word limit) [example reports IED #72 - 79].

The overall structure of OIOS-IED reports should follow two basic principles:

• Deductive logic - the report "tells a story" with the evaluation results. It does not simply present results against a pre-set list of questions; and

• Logical flow - the report is easy for an outside reader to understand, clearly

written with a logical flow within and among sentences and paragraphs.

EXECUTIVE SUMMARY

The Executive Summary is of special importance because it provides a brief synopsis (usually two to three pages) of the evaluation. This makes it possible for busy Member State delegates, UN Secretariat programme managers and others to get a good idea of what the evaluation is about without reading the entire report. The executive summary encapsulates the overall story of the evaluation report in a clear, concise and compelling way. It includes the following sections:

- A description of the evaluation, including why it was conducted;
- A brief discussion (typically no more than two sentences) of the evaluation methodology;
- A summary of major results; and
- A summary of conclusions and recommendations (in bullet form).

INTRODUCTION

The introduction makes a clear statement of the genesis of the evaluation, including a reference to the OIOS-IED Strategic Risk Assessment Framework (Part I, Section 2.1), where applicable. It contains an explicit statement of the evaluation objective and discusses scoping decisions made. It also makes a brief statement of how the evaluand was engaged throughout the evaluation in order to ensure utilization of the evaluation report.

The introduction chapter sets the stage. It is not sufficient to refer to the inception paper (Part II, Section 2.9) for further information. The following elements in Table 18 should typically be addressed.

Table 18: Elements to Address in the Introduction of Evaluation Reports

Subject and scope of the evaluation	Short description of the subject and focus of the evaluation: • Key focus of the evaluation (subject); and • Information about the time period, geographical area, expected results, and so forth, covered (or not covered) by the evaluation and reasons for the choice (scope).
Purpose and objective of the evaluation	Generally speaking, evaluations are conducted for the purpose of accountability, learning and/or decision-taking. The purpose explains why an evaluation is conducted: • Who requested it and their reasons for doing so (reference the OIOS-IED Strategic Risk Assessment Framework (Part I, Section 2.1), and the relevant GA resolution if a GA-mandated report); and • How evaluation results will be used. The objective explains what the evaluation attempts to achieve.
Key evaluation criteria and questions	Information about the evaluation criteria assessed and (if word limit permits) a summary of the key evaluation questions (Part II, Section 2.5).
Evaluation period	The timing and duration of the evaluation activities should be indicated.
Structure of evaluation report	Ideally (if word limit permits), a short paragraph introducing the chapters of the evaluation report and their contents - e.g., "The report has been organized as follows: …"

METHODOLOGY

The methodology section of the evaluation report explains the evaluation methods used to arrive at the results. It discusses the evaluation design (Part II, Step 2), data collection (Part II, Step 3) and data analysis (Part II, Step 4) methods as well as data sources. It refers to efforts to be gender equality and human rights-responsive (Part I, Section 1.6), and explains the specific sampling strategies (Part II, Section 3.3) used to ensure adequate coverage and representativeness and to avoid bias, where relevant.

The methodology section also discusses any methodological or practical limitations or challenges and, where applicable, any creative or innovative approach-

es used to tackle these. It also mentions the use of consultants (Part I, Section 2.3). The following elements in Table 19 should typically be addressed.

Table 19: Elements to Address in the Methodology Section of Evaluation Reports

Data sources and data collection	Explains and provides the rationale for the selected methodology, including sampling. Answers who/what provided data (e.g., documents, project staff, external stakeholders, beneficiaries) and how data was collected (e.g., through interviews, focus groups, documentary review, surveys, direct observation). Describes any gender equality and human rights considerations in the design of the data collection process (e.g., gender-balanced selection of interviewees).
Data analysis	Explains methods applied and steps taken to compile, analyze and triangulate data in order to identify key evidence and arrive at evaluation results.
Limitations and validity of evaluation results	Describes key limitations to the evaluation process (e.g., in terms of time and expertise, availability of stakeholders and beneficiaries, survey response rates, security situation, quality of results-based management) and - given those limitations - makes a frank assessment of the validity and reliability of data and evaluation results.
Use of advisory panels and subject-matter experts	Includes reference to subject-matter experts and advisory panels, their composition and roles where applicable.

BACKGROUND

A background section spells out the most important information to convey about the evaluand, including its PIP or TIP (Part II, Section 2.3). It provides factual information for a better understanding of the context of the subject of the evaluation - e.g., on national goals, global statistics, complementary interventions, experiences elsewhere, related scientific insights. It can also include information on programme duration, location(s), key stakeholders and expected beneficiaries, budget, expenditures, funding availability, intervention logic, previous monitoring and evaluation findings.

RESULTS

The results section forms the core of OIOS-IED reports. Its structure reflects the analytical framework as expressed in the evaluation design (Part II, Step 2) and in particular the underlying PIP/TIP. Results capture the "bottom line" in a descriptive but succinct way. They flow logically from supporting evidence with sound analysis for major assertions. Together, the results and supporting evidence provide the answers to the questions implied by the evaluation objectives. They provide adequate information on gender equality and human rights aspects. OIOS-IED evaluation results should be:

- Organized - there is an internal coherence to the architecture of the results. They convey, at a glance, the main "story" that results from the inquiry and weave it together as a tight, integrated whole;

- Identifiable - result statements are clearly identifiable and are not "lost" somewhere in the text. They are typically labelled as A, B, C, and so forth;

- Relevant - results relate clearly and directly to the evaluation questions. The supporting detail has a logical, sensible relationship to the issue being addressed;

- Substantive - results provide decision-makers with compelling information;

- Precise - results statements accurately and succinctly state the main results of the evaluation. They are unambiguous. The text is free of extraneous material - i.e., information that is not central to the results;

Figure 12: Sequence of Steps for Vetting GA Reports
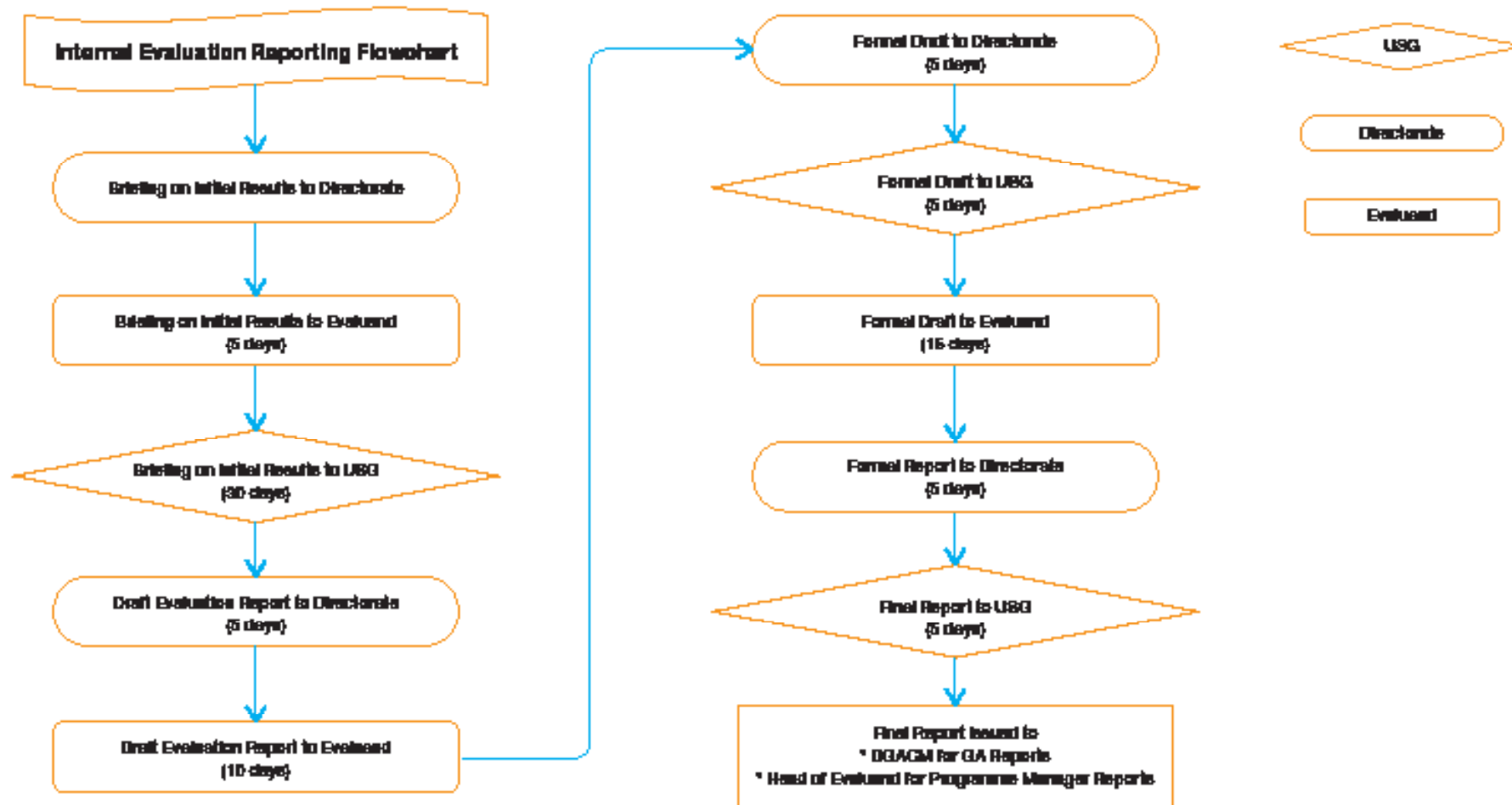
- Persuasive - results are supported by sufficient evidence to convince the reader of their validity;

- Analytical - results not only describe what is, but also explain why things happen or do not happen; and

- Objective - results statements are objective and do not reflect and are not mingled with the views of the evaluators.

Evaluation credibility relies on solid evidence. However, the presentation of evidence is one of the most challenging tasks of an evaluator. The challenge is to present sufficient but not excessive evidence. The following principles help address these challenges:

- Distinguish between background information required to understand a result and evidence supporting a result. Consider moving background information to the background chapter;

- Balance the presentation of evidence (quantitative and qualitative) and rely on different data sources;

- Identify and present key evidence only;

- Aggregate evidence as much as possible (e.g., aggregated survey results or examples); and

- Use text and visuals to present supporting evidence.

Using text for presenting supporting evidence

Evidence from surveys or interviews can be presented in text form (Example 1). When using a magnitude qualifier such as "most" or "some" for survey and interview data, it is necessary to have the quantitative data to support it. Generally, the term "most" or "a majority" should only be used for percentages of greater than 50 percent. Also, when summarizing data from a small sample or universe, raw numbers should be cited (e.g., six of 15 Security Council members). When presenting evidence from interviews, it is particularly important to distinguish between majority and minority views. Single voices can help to articulate key issues, but are never sufficient evidence by themselves. Similarly, quotes should be used very cautiously (if at all) and can only serve to illustrate a widely shared view.

Example 1: "Almost all stakeholders consulted for this evaluation continue to see the GenCap project as highly relevant for improving humanitarian response. In the surveys, 96 per cent of the external stakeholders who expressed an opinion on the subject recommended either maintaining (58 per cent) or expanding (38 per cent) the project. Among advisers, support was similarly overwhelming, with 94% of those expressing an opinion supporting maintenance (35 per cent) or expansion (59 per cent) of the project. … Interviews at both the global and the country level produced similar results. The vast majority of interviewees strongly supported the continued existence of GenCap. Only a very small, but vocal, minority suggested closing or fundamentally changing the project. Most of the time, the debate thus focused not on whether GenCap could be useful in the future, but on how deployments could be fine tuned to add the most value. …"

Using visuals for presenting supporting evidence

Visual information is used to enhance the appeal and effects of OIOS-IED evaluation reports, and particularly the results chapter. Tables (text or numeric), graphs and charts as well as illustrations are all used to convey information in a visual way. Used properly, visual additions help convey messages more effectively, add interest for the reader, break up the monotony of continuous text and help the reader focus on key points of interest. However, misuse of these tools can have the opposite effect. When deciding how to use visuals, it is important to keep in mind the report's audience. To assist in using visuals effectively, a few basic guidelines are helpful. Overall, they should be:

- Simple and easy to understand and avoid elaborate presentations;
- Used for information that can easily be communicated without text (text should not repeat what a visual is conveying and vice versa);
- Clearly labelled;
- Easily distinguished and understood;
- Culturally appropriate;
- Well placed within the report;
- Consistent with numbers and labels; and
- Appropriately referenced.

Tables are best used to present numerical information or to organize data along a set of criteria (Example 2). They can also be used to summarize text-based information under particular categories. Often they provide the basis for other forms of presentation such as graphs or charts. It might be more appropriate to place tables in an annex and to use a chart in the main report to summarize the data. Tables should be accompanied by brief explanations in the narrative about what they are meant to show and how they are to be read. There are a few points about the design and use of tables to remember:

- Make tables simple and accurate; when selecting a format, do not use too many lines, columns or rows; always be certain about the number entered;

- Clearly label rows and columns and try to avoid using abbreviations;
- When showing percentages, round off to the nearest whole number - do not use decimal places;
- Always show the total number (N =) for rows and columns;
- If appropriate, provide sums and averages for each cell so that readers can easily make comparisons; and
- Identify data sources.

Example 2: "Over the programme period, the number of countries where Computer-Based Training (CBT) centres with AML/CFT training modules were set up increased from 4 to 46, with sometimes more than one centre per country. … As can be seen from Table 5, the largest number of countries with newly-established centres was in Africa, and the greatest increase was in 2006."

Table 5.   Number of CBT training centres

Year Established	Africa	Asia and Pacific	Caribbean	Eastern Europe	Latin America	Grand Total
2002		3				3
2004	1					1
2005	5	4				9
2006		5	10		4	19
2007	8				1	9
2008	3			1		4
2009	1					1
Grand Total	18	12	10	1	5	46

Graphs and charts are another effective way to communicate key points in evaluation reports, usually with little supporting text (Example 3). There are a few points to consider when using graphs and charts:

- Use both upper and lower case letters, and only one or two type fonts;
- Avoid busy patterns;
- Use colours and patterns that print well on black and white devices;
- Keep as simple as possible;
- Keep scales honest;
- Use titles and sub-titles to convey messages;
- Identify data sources; and
- Place supporting data in annexes.

Example 3: "Survey respondents rated the comprehensiveness, reliability and policy relevance of both the WESP (World Economic Situation and Prospects] and the WESS [World Economic and Social Survey) higher than some of their other characteristics."

The main point to consider when using illustrations (e.g., sketches, maps, photographs) is that to be effective they must be relevant to the topic. They should be used for a specific purpose and to communicate a particular point - e.g., to show progress over time. They should not be used for decoration. Illustrations require explanations in the narrative.

Alternative ways to present evidence

In addition to written evaluation reports, alternative ways of presenting evidence, such as videos, are increasingly used by evaluators.

Table 5: Characteristics of WESP and WESS as perceived by DESA stakeholders who had read them
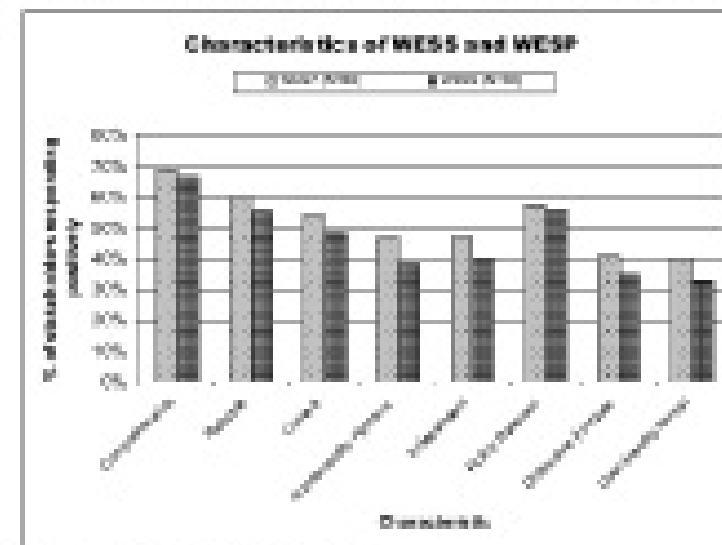
Source: OIOS stakeholder survey

CONCLUSIONS

Conclusions flow logically from the results. They answer the "so what" question of the report that warrants the attention of decision-makers. Conclusions interpret evaluation results. They correspond to a doctor's diagnosis or a judge's verdict.

The conclusions chapter should provide the evaluators' professional views on the assessed criteria and significant issues. The conclusions build the bridge between the past and present (results) with the future (recommendations).

Sound conclusions:

- Answer key evaluation questions and focus on significant issues;
- Reflect the evaluators' professional views and judgment on key evaluation results;
- Add value by addressing future key opportunities and challenges;
- Do not just repeat or summarize results; and
- Do not introduce any new information.

A helpful way to distinguish evidence, evaluation results and conclusions is the degree of interpretation and judgement required by the evaluator. This is zero for evidence, medium for results and high for conclusions and recommendations (Figure 13).

RECOMMENDATIONS
Where considered necessary, recommendations are formulated that set out actions to respond to problems or opportunities identified in the report. In 2011, OIOS introduced a new system for categorizing recommendations into three groups: critical recommendations, important recommendations and opportunities for improvement.

Critical recommendations
Critical recommendations are those that address significant and/or pervasive deficiencies in governance, risk management or internal control processes, such that reasonable assurance cannot be provided regarding the achievement of programme objectives. Any critical recommendations rejected by the evaluand may be elevated to the UN Deputy Secretary-General (DSG) or Secretary-General (S-G) if necessary. OIOS follows up on implementation of critical recommendations on a quarterly basis.

Figure 13: Whose Voice?

King Zollinger, 2012
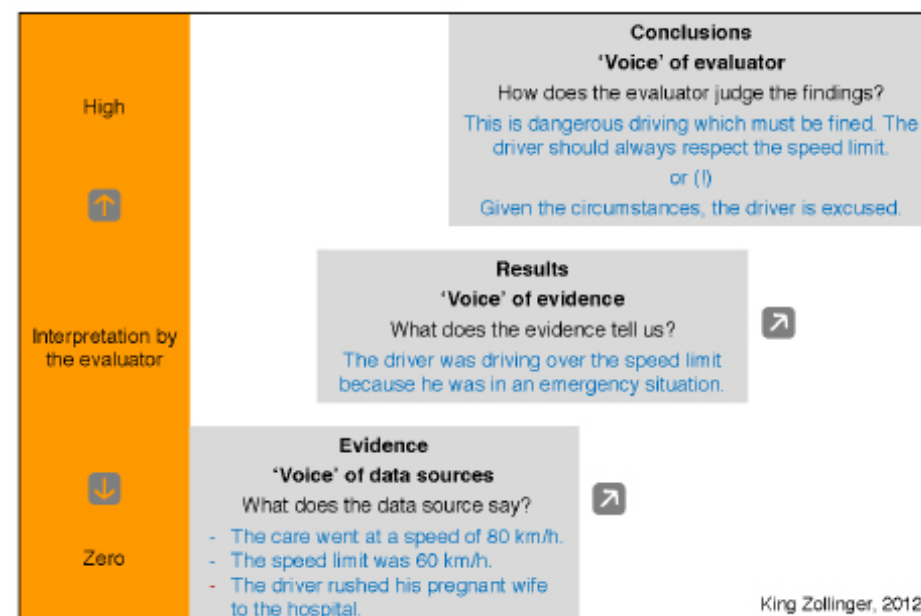
Important recommendations
Important recommendations address reportable deficiencies or weaknesses in governance, risk management or internal control processes, such that reasonable assurance might be at risk regarding the achievement of programme objectives. Important recommendations are followed up on an annual basis.

Opportunities for improvement
Opportunities for improvement are suggestions that do not meet the criteria of either critical or important recommendations, and are only followed up as appropriate during subsequent oversight activities.

The number of recommendations should be kept to a minimum (typically 5-12). To increase the ownership and utility of evaluation reports, evaluands are consulted in the formulation of the recommendations. The basic characteristics of OIOS-IED recommendations are that they be:

- Relevant - they are clearly based on and explicitly linked to evaluation results;
- Prioritized - they are ranked in order of importance or urgency ("critical", "important" and "opportunities for improvement");
- Targeted - they address the appropriate body/ies;
- Time-bound - they specify by when recommendations should be implemented;
- Clear - they are as specific as possible, while avoiding excessive prescriptiveness;
- Feasible - they are capable of being accomplished within the timeframe and resources available; and
- Strategic - they have the potential to bring about real change.

Many factors have to be considered when preparing recommendations [📖 BetterEvaluation]. Some recommendations relate to probity, transparency, accountability, propriety and so on and are not directly related to the programme-specific PIP or TIP. However, where recommendations are made about other aspects of programme implementation and outcomes, then reference to the PIP/TIP can provide a useful litmus test for determining the validity and relative importance of recommendations. One important consideration is that recommendations should be defensible in terms of being able to demonstrate through "if-then" reasoning - i.e., that when implemented they will contribute to improved outcomes and impacts. Referring to the PIP/TIP (Part II, Section 2.3) can provide a framework for this if-then reasoning and prevent excessive preoccupation with minor programme improvements.

ANNEXES

Annexes provide OIOS-IED evaluators with a means of providing additional evidence or more detailed information on one or more of the evaluation issues while allowing the main body of the text to be kept relatively lean.

Annexes are generally optional inclusions in OIOS-IED reports. By contrast, the one item that must be annexed to every OIOS-IED report, for both GA-mandated evaluation reports and non-GA reports, is the full text of the memorandum containing the formal comments provided by the evaluand (Part II, Section 5.1). The PIP/TIP used in the evaluation must also be annexed to every report.

Table 20 lists examples of annexes that may be included in OIOS-IED evaluation reports.

Table 20: Examples of Annexes

	Annexes
Methodological supplements	Evaluation methodology matrices; sampling strategies; interview questions; questionnaires; and so forth.
Itinerary	Itinerary of any field visits, including project sites visited.
List of individuals met	In order to ensure confidentiality, OIOS-IED does not include names of individuals interviewed in the report but rather provides the names of institutions or organizations they represent, if that is appropriate.
Documents consulted	Including websites.
Background information	Additional contextual information to supplement the background chapter - e.g., maps, if necessary.
Supplementary evidence	Evidence and analyses presented in more detail than in the main text - e.g., survey results and case studies.

Step 6:   Dissemination

| Step 1: Announcement | Step 2: Evaluation Design | Step 3: Data Collection | Step 4: Data Analysis | Step 5: Report Preparation | Step 6: Dissemination | Step 7: Housekeeping | Step 8: Tracking Follow-Up |

In keeping with UNEG guidelines [➜ UNEG. 2010c], all OIOS-IED evaluations have a plan for report dissemination and follow-up. Some elements of this plan are dictated by standard operating procedure, especially in the case of evaluation reports mandated by the GA. Other elements are determined on a case-by-case basis. Wherever possible, the OIOS-IED evaluation team should think through its dissemination strategy already during the evaluation design stage (Part II, Step 2), and state as much of this plan as is known at this early stage of the evaluation in the inception paper (Part II, Section 2.9).

Aspects of report dissemination that are followed as standard operating procedure are as follows:

- As a rule, all final reports, GA reports and non-GA reports alike, are placed on the OIOS-IED intranet. They are also shared with the Office of the OIOS Under-Secretary-General (OUSG), OIOS's Internal Audit Division (IAD) and Investigations Division (ID), the JIU and BOA, and the evaluand by way of an official memorandum, which, in addition to standard language introducing the report, conveys to the evaluand the requirement that a management response and recommendation action plan (Part II, Step 8) be completed within a specified time frame. For all reports, management responses are appended to the final report.

- GA reports are public documents, and as such are placed both on the OIOS intranet and internet [📖 OIOS] sites.

- To help promote utilization, OIOS-IED engages with evaluands at least once after the evaluation is completed to further discuss the evaluation - typically providing more data that did not make it into the report - and providing advice and consultation on how to address issues raised in the evaluation as well as how to implement the recommendations.

Special procedures apply to the presentation of GA reports submitted to intergovernmental bodies, including the GA Fifth Committee and the CPC, namely:

- The final report is sent to DGACM for editing, formatting and translation. During its processing of the report, DGACM might revert to the OIOS-IED Section Chief with questions for clarification. Reports submitted to DGACM are assigned a "slot date" by which the report must be submitted or it is considered late;

- The Team Leader ensures that the report is fully annotated (Part II, Section 7.1) with evidence for all of its results, conclusions and recommendations.

- The OIOS-IED evaluation team meets to review possible questions from, and answers to, Member States;

- The evaluation team prepares an introductory statement [IED #27] for the OIOS USG (or other senior OIOS staff member) to introduce the report;

- The team compiles a binder for the CPC or Fifth Committee session(s), including the final evaluation report, the annotated report, possible questions and answers, relevant GA resolutions, and key data analysis summaries.

- During Fifth Committee and CPC sessions, one or more OIOS-IED staff members are assigned minute-taking responsibilities. OIOS-IED is responsible for drafting the CPC report sections on the discussions of its reports [IED #80]. These are typically due to the CPC Secretariat one day after conclusion of the formal session for the report.

In addition to these required steps, there are optional steps that OIOS-IED evaluation teams might wish to consider in order to maximize the utilization of their report. These include the following:

- Should anything arise during the evaluation that OIOS-IED chooses not to include in its report, either out of confidentiality concerns or other sensitivities, OIOS-IED might wish to draft a confidential memorandum, separate from issuance of the final report, informing the evaluand's USG of these other issues;

- The evaluation team might consider drafting an article for inclusion in OIOS-IED's newsletter, providing a brief summary of the evaluation, its major results and recommendations, and any noteworthy early outcomes;

- The evaluation team might consider designing a brief fact sheet on the evaluation, written in more accessible layman's language than official reporting conventions allow, for use both by OIOS-IED and/or the evaluand to promote the evaluation;

- Throughout the evaluation, evaluation teams should determine the utility of encouraging the evaluand to share the report with its governing body (and even suggest a presentation by OIOS-IED to the governing body) as a means of enhancing its report's utilization. If evaluation teams choose this route for a GA report, it is important that OIOS-IED not directly undertake any dissemination activities with the evaluand's governing body before OIOS-IED has briefed the CPC, Fifth Committee or other GA body.

- If case studies have been undertaken as part of the evaluation, the evaluation team might wish to repackage the key takeaways and share these with the evaluand in the form of a briefing note, fact sheet, aide mémoire, project summary [IED #31, #31a], or other format.

There are numerous external resources aimed at arming evaluation teams with strategies for enhancing the utilization of their evaluations [➔ Quinn Patton. 2008a] [📖 BetterEvaluation].

Step 7:   Post-Evaluation "Housekeeping"

| Step 1: Announcement | Step 2: Evaluation Design | Step 3: Data Collection | Step 4: Data Analysis | Step 5: Report Preparation | Step 6: Dissemination | Step 7: Housekeeping | Step 8: Tracking Follow-Up |

Upon delivery of the final evaluation report to the DGACM, OIOS-IED undertakes three main activities before it begins follow-up on report recommendations later on.

The following main sections discuss expectations around each of these activities.

- Report Annotation (Part II, Section 7.1)
- File Management (Part II, Section 7.2)
- Lesson Learning Sessions (Part II, Section 7.3)

7.1   Report Annotation

At the latest once an OIOS-IED evaluation report has been finalized, the evaluation team produces an annotated report [IED #28] as a matter of good evaluation practice and a means of ensuring that its evidence is readily available when interacting with key stakeholders after dissemination of the evaluation report. There is no standard approach to the timing of report annotation. Some evaluators prefer to annotate their report in "real time" – i.e., while drafting the report, then strip away the annotations before submitting it for comment by their respective Section Chief, the OIOS-IED Directorate, the OIOS USG and the evaluand. Others prefer to annotate the report after the report is finalized.

Report annotation typically involves footnoting the data sources underlying each results statement and conclusion in the report, the goal being that any cold reader could follow the annotation and be confident that ample evidence exists. For example, if a results statement is made based on certain survey responses, the questionnaire, question number(s) and responses should be referenced to support that statement.

Report annotation puts the evidence base underlying its claims at OIOS-IED's immediate disposal, helping it to respond to questions during CPC or GA Fifth Committee sessions and other interactions with Member States, and to have a constructive dialogue with the evaluand.

Annotation is particularly important because it is often not the Team Leader who engages in such interactions, but rather his or her Section Chief, OIOS-IED's Director or Deputy Director, or the OIOS USG.

7.2   File Management

In 2008, OIOS-IED developed an internal file management structure to ensure that all information in the Division is maintained in a consistent and efficient manner. All materials critical to evaluations should be saved in a clearly identifiable manner. This ensures that the evaluation team can access relevant documents easily should the need arise, and helps future evaluation teams do so as well.

At minimum, evaluation Team Leaders should ensure that the following documents are saved on the OIOS-IED server:

- Evaluation notification memo (Part II, Section 1.1);
- Team compact (Part II, Section 2.2);
- Inception paper and any associated annexes (Part II, Section 2.9);
- Evaluation budget;
- All data collection instruments (Part II, Section 2.9);
- All raw data gathered through various data collection methods (e.g., databases, interview notes);

- All data analysis summaries (e.g., tabulated survey questionnaires);
- All drafts of the evaluation report;
- All quality assurance checklists (Part I, Chapter 3);
- CVs for all consultants;
- All consultant contracts and other types of contracts;
- All critical correspondence with the evaluand and others; and
- Management response and recommendation action plan (Part II, Step 8)

7.3   Lesson Learning Sessions

OIOS-IED has a well-developed learning mechanism. At the conclusion of each inspection and evaluation (Part I, Section 1.4), the OIOS-IED evaluation team, together with its Section Chief, conducts a lesson learning session to discuss what went well and what went less well in the conduct of the project. Other staff members of OIOS-IED are invited to these sessions.

The sessions should be rooted in how well the team has succeeded in achieving the results targeted in the OIOS-IED PIP (Part I, Section 1.2), and why. During the sessions, teams should consider framing their discussion around the following questions:

- Did we meet our deadlines? Why or why not? If not, how much slippage did we experience at key junctures?

- Did we stay within our budget? Why or why not?

- To what extent did we adhere to the OIOS-IED Quality Assurance System (QAS) (Part I, Chapter 3)? Why or why not?

- To what extent did we pursue the most relevant lines of inquiry? Why or why not?

- To what extent did we produce the most credible analysis possible? How would we rate the quality of our evaluation report? Why?

- How effective were we in engaging key stakeholders so that the evaluation report, its conclusions and recommendations, would be utilized? Why?

- How efficient were we in getting the job done? Why?

- How clear was our internal communication with each other, and our external communication with others - e.g., the Office of the OIOS Under-Secretary-General (OUSG) and the evaluand?

- How clearly delineated were individual team members' roles and responsibilities, as per the team compact (Part I, Section 2.2)? How successfully did we adhere to these agreed roles and responsibilities?

- How clear and helpful was the guidance provided by the Chief of Section and the OIOS-IED Directorate?

- How successfully did we engage consultants (Part I, Section 2.3) in order to benefit from their expertise?

- How impactful was our evaluation (or how impactful is it likely to be) in terms of helping to improve the UN Secretariat programme, as far as we know? Why?

- How successful were we in ensuring that our evaluation did not cause undue negative consequences to the programme, and in particular its targeted beneficiaries?

- How effectively did we handle threats to OIOS-IED's independence, if applicable?

- How effectively did we incorporate a gender and human rights (Part I, Section 1.6) as well as environmental sustainability lens into our conduct of the evaluation?

- How professionally did we comport ourselves as evaluators - e.g., exhibiting integrity and independence, avoiding conflicts of interest and flagging them when they occurred?

- How systematically did we explore the impact of OIOS-IED's previous evaluations of the evaluand?

OIOS-IED evaluation teams document lessons learned by answering these questions in a summary document [IED #29], shared with all OIOS-IED staff and stored on the OIOS-IED shared drive for future reference. As a learning and improving division, OIOS-IED synthesizes these lessons learned summaries every year. This annual recapitulation of key success stories and shortcomings helps reflect on what it is doing so that it can undertake concrete actions to capitalize on its good practice and avoid future pitfalls.

Step 8:   Tracking Follow-Up to Recommendations

Step 1: Announcement	Step 2: Evaluation Design	Step 3: Data Collection	Step 4: Data Analysis	Step 5: Report Preparation	Step 6: Dissemination	Step 7: Housekeeping	Step 8: Tracking Follow-Up

Beginning in 2011, OIOS introduced a new system for categorizing recommendations into three groups - i.e., critical recommendations, important recommendations and opportunities for improvement.

Critical recommendations are those that address significant and/or pervasive deficiencies in governance, risk management or internal control processes, such that reasonable assurance cannot be provided regarding the achievement of programme objectives. Those rejected by the evaluand may be elevated to UN Deputy Secretary-General (DSG) or S-G if necessary. Critical recommendations are followed up on a quarterly basis. Important recommendations are those that address reportable deficiencies or weaknesses in governance, risk management or internal control processes, such that reasonable assurance might be at risk regarding the achievement of programme objectives. Important recommendations are followed up on an annual basis. Opportunities for improvement are suggestions that do not meet the criteria of either critical or important recommendations and are only followed up as appropriate during subsequent oversight activities.

Once a report is sent to the evaluand as a final report, the evaluand is asked within the first month to submit a recommendation action plan [IED #70] for implementation of inspection or evaluation recommendations. The action plan includes the actions to be taken, the entity responsible for undertaking it and target dates for completion.

For evaluation reports submitted to the CPC, which may change OIOS-IED recommendations or introduce its own, this process might be delayed until after the CPC completes its deliberations (every other June according to the CPC schedule).

The system used to track recommendations is called Issue Track, which is a database developed to integrate the recommendation databases of all three OIOS divisions into a single departmental system.

For OIOS-IED, the administrative support person of each OIOS-IED section enters the recommendations into Issue Track once the report is finalized. All data entry into Issue Track is done through a screen called Recommendation Form (RF). Data for the RF must be signed off by the Team Leader, followed by the Section Chief.

During the tracking process, evaluand responses on the status of recommendations are received by the designated OIOS-IED Issue Track focal point and entered into Issue Track by the Team Leader. If a Team Leader no longer works for OIOS-IED, then the Section Chief assigns another person to be responsible for Issue Track for that project, usually either another team member or the Section Chief him/herself. The Team Leader ensures completeness of the responses and follows up directly on non-responses, with questions or requests for supporting documentation. Responses are typically entered in Issue Track within one week of their receipt. When complete, the Office of the OIOS Under-Secretary-General (OUSG) consolidates the data and prepares statistics for OIOS annual and semi-annual reports that are presented to the GA Fifth Committee. OIOS-IED's Director is ultimately held responsible for any errors in the recommendations data at this stage of the process.

It is ultimately up to the judgment of the Team Leader, and final approval of the Section Chief, to determine whether a recommendation has been implemented. The following evidentiary standards should be applied in making this determination:

- The original intent of the recommendation is satisfied;
- All relevant documents have been produced;
- All relevant meetings have been conducted;
- Evidence of change in work procedures is obtained; and
- Evidence of change in behaviours is obtained;

Team Leaders rely on various data collection methods as means of verification, including document reviews of the evidence provided, interviews, surveys, website reviews and direct observation.

The OIOS-IED Quality Assurance System (QAS) (Part I, Chapter 3) includes a checklist for report follow-up [IED #82]. This checklist helps evaluation teams ensure that critical items described in this section are included in the report drafting process.

With regard to OIOS-IED's commitment to monitoring its own progress in achieving the impact targeted in its PIP (Part I, Section 1.2), OIOS-IED's Triennial Reviews (Part I, Section 1.4) report on the evaluands' implementation of recommendations in those evaluations that have been submitted to the GA. In addition, beginning in the 2014-2015 evaluation cycle, all OIOS-IED evaluations began incorporating into their methodology a brief line of inquiry to follow up, wherever applicable, on the impact of OIOS-IED's previous evaluations.

Acronyms and Abbreviations

The following acronyms and abbreviations are those cited in the manual. There are many more such acronyms and abbreviations of immediate relevance to OIOS-IED in its day-to-day work. The UN maintains a more comprehensive list of key acronyms and abbreviations, as well as key terms used in the UN system.

BOA	United Nations Board of Auditors
CEB	Chief Executives Board
CPC	Committee for Programme and Coordination
C34	Special Committee on Peacekeeping Operations
DESA	Department of Economic and Social Affairs
DFS	Department of Field Support
DM	Department of Management
DPA	Department of Political Affairs
DGACM	Department for General Assembly and Conference Management
DPKO	Department of Peacekeeping Operations
DSG	Deputy Secretary-General
GA	United Nations General Assembly
IAAC	Independent Audit Advisory Committee
IAD	Internal Audit Division
ID	Investigations Division
IED	Inspection and Evaluation Division
IMDIS	Integrated Monitoring and Documentation Information System
JIU	Joint Inspection Unit of the UN System
M&E	Monitoring and evaluation
MECD	Monitoring, Evaluation and Consulting Division
OCHA	Office for the Coordination of Humanitarian Affairs
OHCHR	Office of the High Commissioner for Human Rights
OHRM	Office of Human Resources Management
OIOS	Office of Internal Oversight Services

OLA	Office of Legal Affairs
OLS	Ordinary Least-Squares (OLS) regression
OPPBA	Office of Programme Planning, Budget and Accounts
OUSG	Office of the Under-Secretary-General
PIP	Programme Impact Pathway
PKO	Peacekeeping operation
PPBME	Regulations and Rules Governing Programme Planning, the Programme Aspects of the Budget, the Monitoring of Implementation and the Methods of Evaluation
QAS	Quality Assurance System
QDAS	Qualitative Data Analysis Software
QSA	Support account (for peacekeeping-related work)
RF	OIOS Issue Track Recommendation Form
S-G	Secretary-General
SPM	Special political mission
ST/AI	Administrative instruction
ST/SGB	UN Secretary-General's bulletin
TIP	Thematic Impact Pathway
ToR	Terms of reference
UN	United Nations
UNCTAD	United Nations Conference on Trade and Development
UNDAF	United Nations Development Assistance Framework
UNEP	United Nations Environment Programme
UNEG	United Nations Evaluation Group
UN-HABITAT	United Nations Human Settlements Programme
UNHCR	United Nations High Commissioner for Refugees
UNODC	United Nations Office on Drugs and Crime
USG	Under-Secretary-General
VTC	Video teleconference

Glossary

Main sources:

[IED #3]

[➔OECD DAC. 2009a]

[➔ Trinidad. 2008a]

Accuracy

Data accuracy is the extent to which data were obtained in a clear and well-defined manner from different data-sets across cases, space and time as well as the extent to which they truthfully reflect facts and figures such as dates, percentages and numbers of persons interviewed.

Activities

Action taken to transform inputs into outputs.

Anonymity

Evaluator action to ensure that the identity of subjects cannot be ascertained during the course of an evaluation, in evaluation reports, or in any other way.

Assumptions

Describes factors or risks that can affect the success or failure of a project/programme.

Attribution

The establishment of a causal link between observed changes and a specific intervention, and to the actor to be credited for the results observed.

Baseline

Facts about the condition or performance of subjects prior to treatment or intervention.

Benchmark

Reference point or standard against which performance or achievements can reasonably be assessed.

Beneficiaries

The individuals, groups, or organizations, whether targeted or not, that benefit, directly or indirectly, from a development intervention.

Bias

Bias might result in overestimating or underestimating certain characteristics of the population. It might result from incomplete information or invalid data collection methods and might be intentional or unintentional.

Closed-ended question

A question that limits respondents' answers to pre-determined response categories. Multiple choice and yes/no questions are examples of closed-ended questions.

Coding

The process of organizing data into sets of categories to capture the meaning or main themes in the data. Coding is usually done in the analysis of qualitative data, but quantitative data can also be coded.

Conclusions

Conclusions point out the factors of success and failure of the evaluated intervention, with special attention paid to the intended and unintended results and impacts, and more generally to any other strength or weakness. A conclusion draws on data collection and analyses undertaken, through a transparent chain of arguments.

Conflict of interest

Conflict of interest can be personal or organizational. Personal conflict of interest refers to a situation where a person's private interests interfere or might be perceived to interfere with his or her performance of official duties. Organizational conflict of interest refers to a situation where, because of other activities or relationships, an organization is unable to render impartial services and the organization's objectivity in performing mandated work is or might be impaired, or the organization has an unfair competitive advantage.

Contribution

The link between the activities carried out by various organizational units to arrive at final products or services delivered to end-users to accomplish a desired result.

Control group

A group of subjects that have not been exposed to an intervention. The control group should resemble the programme group (the subjects which have been exposed to the intervention), so that systematic differences

between the two groups might be attributed to the effects of the intervention once other plausible alternative hypotheses have been eliminated or discounted. True control groups are formed through randomization.

Convenience sample
A non-random sample drawn from the target population because of ease of access.

Correlation
The extent to which two or more variables are related to each other.

Counterfactual
The situation or condition that hypothetically might prevail for individuals, organizations or groups had the intervention not taken place. Counterfactual analysis, for example, requires a comparison between what actually happened and what would have happened in the absence of the intervention.

Dissemination
The process of communicating information to specific audiences for the purpose of extending knowledge and with a view to modifying policies and practices.

Efficiency
A measure of how well inputs (funds, staff, time, and so forth) are converted into outputs.

Effectiveness
The extent to which a development intervention's objectives were achieved, or are expected to be achieved, taking into account their relative importance.

Evaluability
Extent to which an activity or a programme can be evaluated in a reliable and credible fashion.

Evaluand
An entity that is subject to an inspection or evaluation.

Evaluation
A process that seeks to determine as systematically and objectively as possible the relevance, effectiveness and impact of an activity in the light of its goals, objectives and accomplishments.

Evaluation criteria
The characteristics against which a programme or project is evaluated.

Evaluation design
Used broadly, this term describes the complete approach and plan for evaluation process. Used more narrowly, it refers to a specific strategy for answering specific evaluation questions.

Evaluation questions
The questions that will be answered through data collection, analysis and interpretation. Evaluation questions are developed from the evaluation goals and objectives and state specific information needs. They focus on aspects and outcomes of the project.

Evaluator
An individual involved in all stages of the evaluation process, from defining the terms of reference and collecting and analysing data to developing findings and making recommendations. The evaluator might also be involved in taking corrective action or making improvements.

Evidence
The information presented to support a finding, an assertion or a conclusion that is sufficient, competent, and relevant.

Experimental design
The classic experimental design includes treatment and control groups that are studied before and after an intervention. Subjects are chosen at random and must have shared characteristics. The treatment group is exposed to the intervention and the control group is not. After the intervention, a comparison is made between the two groups to determine if there are any significant changes between the two groups. If it is determined that the group receiving the treatment has had a more significant improvement than the control group, it can be concluded that the intervention has been successful.

Focus group
A group selected for its relevance to an evaluation that is engaged by a trained facilitator in a series of discussions designed for sharing insights, ideas, and observations on a topic of concern to the evaluation.

Hawthorne effect
The term "Hawthorne effect" is used to explain situations where an experiment cannot be trusted because the very fact that the experiment is taking place is influencing the results obtained. This suggests that programme staff and beneficiaries might behave quite differently from their normal patterns if they know that they are being observed.

Impact
In general, an expression of the changes produced in a situation as the result of an activity that has been undertaken. It includes positive and negative, primary and secondary long-term effects produced by an intervention, directly or indirectly, intended or unintended. Impact might also refer to the ultimate, highest level, or end outcome of an activity or set of activities.

Impact evaluation
Impact evaluation is the systematic identification of the effects - positive or negative, intended or not - on individual households, institutions and the environment caused by a given development activity such as a development programme or project.

Independence
The freedom from conditions that threaten the ability of the oversight activity to carry out its oversight responsibilities in an unbiased manner.

Indicators of achievement
Measure of whether and/or the extent to which objectives and/or expected accomplishments have been achieved. Indicators correspond either directly or indirectly to the objective or the expected accomplishment for which they are used to measure performance.

Inputs
Personnel and other resources necessary for producing outputs and achieving accomplishments.

Intervention logic
A reasoned description of how a programme is expected to attain its objectives using hypothetical cause-effect linkages to show the chain of expected effects between inputs, activities, outputs, outcomes and, ultimately, impact.

Lessons learned
The knowledge or understanding gained from the implementation of a programme, subprogramme or project that is likely to be helpful in modifying and improving programme performance in the future. This knowledge is intentionally collected with the purpose of using it in the future and it includes both positive and negative lessons.

Knowledge management
The systematic creation, organization, storage and sharing of knowledge in order to better achieve organizational goals. A knowledge management strategy might include, inter alia, a description of how the organization learns from projects and makes that learning accessible to people in other parts of the organization.

Mixed-method evaluation
An evaluation for which the design includes the use of both quantitative and qualitative methods for data collection and data analysis.

Monitoring
Monitoring is a continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing development intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds.

Non-response analysis
Non-response analysis is a comparison of respondents' demographic attributes against those of non-respondents, to the extent any of these are known.

Objectives
An overall desired achievement involving a process of change and aimed at meeting certain needs of identified end-users within a given period of time. Objectives should be, to the greatest extent possible, concrete and

time-limited. Achievement of the objectives should be verifiable either directly or through evaluation. Indicators of achievement should be given where possible.

Objectivity
An unbiased approach that requires evaluators to perform engagements in such a manner that they believe in their work product and that no quality compromises are made. Objectivity requires that evaluators do not subordinate their judgment on oversight matters to others.

Observation
A data collection method in which the researcher watches and records events and processes.

Outcomes
The likely or achieved short-term and medium-term effects of an intervention's outputs. They can be either intended or unintended, and desired (positive) or unwanted (negative).

Outputs
The final products or services delivered by a programme or subprogramme to end-users, such as reports, publications, training, servicing of meetings, or advisory, editorial, translation or security services, which an activity is expected to produce in order to achieve its objectives.

Performance indicator
An objective measure of a variable that provides a reliable basis for assessing achievement, change or performance. A unit of information measured over time that can help show changes in a specific condition.

Performance measurement
A system for the collection, interpretation of, and reporting for the purpose of objectively measuring how well programmes or projects contribute to the achievement of expected accomplishments and objectives and deliver outputs.

Population / Universe
All persons. documents, events, or other unit of analysis in a given, well-defined group.

Primary data
Information collected directly by the evaluator rather than obtained from secondary sources (data collected by others) to inform an evaluation.

Programme evaluation
A form of evaluation that assesses the overall relevance, efficiency, effectiveness and impact of a UN Secretariat programme or subprogramme and the full range of its inputs, activities, outputs, outcomes and (where applicable) impact; stands in contrast with project evaluation, which looks at a single intervention. Also referred to as "in-depth" evaluation when mandated by the Committee for Programme Coordination.

Programme Impact Pathway (PIP)
A logical framework tool used to identify strategic elements of a project or programme (objectives, inputs and outputs, intended outcomes, activities, indicators) and their causal relationship, as well as the critical assumptions that might influence success and failure.

Purposive sampling
Creating samples by selecting information-rich cases from which one can learn a great deal about issues of central importance to the purpose of the evaluation.

Qualitative methods
The overall data collection and analytical methods for information that is not easily captured in numerical form (although qualitative data can be quantified). Qualitative data typically consist of words and normally describe people's opinions, knowledge, attitudes or behaviours. Examples of qualitative methods are interviews, focus group discussions, direct observations and literature reviews.

Quality assurance
Quality assurance encompasses any activity that is concerned with assessing and improving the merit or the worth of a development intervention or its compliance with given standards.

Quantitative methods
The data collection and analytical methods for information that is expressed or measurable in a numerical form. Quantitative data typically consist of numbers. Examples of quantitative methods are surveys and statistical analyses of available quantitative data.

Quasi-experimental design
Comparison of the relevant state of the world after the intervention with its state beforehand and attribution of any difference to the effects of the intervention. A particular weakness of this design is the possibility that something else besides the intervention accounts for all or part of the observed difference over time.

Random sampling
Drawing a number of items of any sort from a larger group or population so that every individual item has a specified probability of being chosen.

Recommendation
Proposal for action to be taken to enhance the design, allocation of resources, effectiveness, quality, or efficiency of a programme or a project. Recommendations should be substantiated by evaluation findings, linked to conclusions and indicate the parties responsible for implementing the recommended actions.

Relevance
The extent to which an activity, expected accomplishment or strategy is pertinent or significant for achieving the related objective and the extent to which the objective is significant to the problem addressed. Relevance is viewed in the context of the activity's design as well as in the light of the factual situation at the time of evaluation.

Reliability
Consistency or dependability of data and evaluation judgements, with reference to the quality of the instruments, procedures and analyses used to collect and interpret evaluation data.

Sample / Sampling
A sample is a subset of a population, selected through any number of sampling techniques in order to assess attributes of the population. Sampling, which can be probabilistic or non-probabilitistic, is the process of selecting the sample to assess.

Sampling error
The difference between the true results in the population and the estimate of results derived from a sample because the sample studied is not perfectly representative of the population from which it was drawn. In general, sampling error is lower as sample size increases.

Secondary data analysis
A re-analysis of data using the same or other appropriate procedures to verify the accuracy of the results of the initial analysis or for answering different questions.

Self-administered instrument
A questionnaire or report completed by a study participant without the assistance of an interviewer.

Self-evaluation
A form of evaluation undertaken by UN Secretariat programmes of their own performance, primarily for their own use for the purpose of institutional learning.

Stakeholders
Agencies, organizations, groups or individuals who have a direct or indirect interest in the development intervention or its evaluation.

Statistical analysis
A commonly used data analysis technique. Statistical analysis is often used to describe phenomena in a concise and revealing manner. This is known as descriptive statistics. It can also be used to test for relationships among variables or generalize findings to a wider population. This is known as statistical inference.

Stratified random sampling
A sampling strategy that divides the population into subgroups or strata and draws random samples from each stratum.

Structured interview
A type of formal interview that covers a set of specific questions and asks each respondent the same questions, with the same wording, in the same order.

Survey

A widely-used technique for collecting data from a sample drawn from a given population. Surveys are often based on probability sampling, and survey information is usually obtained through structured interviews or self-administered questionnaires.

Sustainability

The probability of continued long-term benefit.

Terms of reference (ToR)

A written document presenting the purpose and scope of the evaluation, the methods to be used, the standard against which performance is to be assessed or analyses are to be conducted, the resources and time allocated, and reporting requirements.

Thematic evaluation

Thematic evaluations typically assess a cross-cutting theme or activity (e.g., implementation of a gender mainstreaming policy or knowledge management) across multiple Secretariat programmes or peacekeeping operations/special political missions. They can also assess the cumulative effects of multiple programmes sharing common objectives and purposes or the effectiveness of coordination and cooperation among programmes.

Timeliness

Evaluations meet their deadlines with no slippage. The information they contain is conveyed to key stakeholders at optimal moments for influencing key decisions.

Triangulation

An attempt to get a fix on a phenomenon or measurement by approaching it via several (three or more) independent routes.

UNEG norms

Seek to ensure that evaluation entities within the UN follow agreed-upon basic principles. They provide a reference for strengthening, professionalizing and improving the quality of evaluation in all entities of the UN system.

UNEG standards

Build upon the UNEG norms. They are drawn from UNEG member good practices, and are intended to guide the establishment of the institutional framework, the management of the evaluation function, and the conduct and use of evaluations. There are 50 standards for evaluation in the UN system. They fall within four broad categories: i) institutional framework and management of the evaluation function; ii) competencies and ethics; iii) conducting evaluations; and iv) reporting.

Unit of analysis

The entity about which data are collected, analyzed and conclusions drawn.

Utility

The extent to which an evaluation produces and disseminates reports that inform relevant audiences and have a beneficial impact on their work.

Validity

The extent to which the data collection strategies and instruments measure what they purport to measure.

Reference Documents and Websites

OIOS-IED Documents, Templates & Examples

Links to all of the following documents can be found on the IED Intranet at https://iseek-newyork.un.org/departmental_page/ied-manual

#1	OIOS Strategic Framework 2014-2015
#2	OIOS Programme Impact Pathway (PIP)
#3	List of key oversight terms used in OIOS activities
#4	In-depth narrative to accompany the IED PIP
#5	OIOS Self-evaluation Policy
#6	Biennial Review: report screening (guidance)
#7	Biennial Review: evaluation reports assessment sheet (template)
#8	Biennial Review: review of evaluation policies (template)
#9	United Nations Secretariat Evaluation Scorecards 2010-2011 (report)
#10	Roles and responsibilities of the various IED staff categories (matrix)
#11	IED evaluation team compact (template)
#12	Advisory panel terms of reference (example)
#13	Memo sharing IED's evaluation work plan for 2014-2015 (example)
#14	Evaluation notification memorandum (example)
#15	Aide memoire attached to notification memorandum
#16	Information brochure (example)
#17	QAS scoping checklist (template)
#18	How to prepare and use Programme Impact Pathways (PIPs) for scoping and reporting an IED programme evaluation (guidance)
#19	IED evaluation design matrix (example)
#20	Stakeholder map (example)
#21	QAS inception paper checklist (template)

#22	Terms of reference for inspections and peacekeeping evaluations (example)
#23	QAS data collection checklist (template)
#24	QAS data analysis checklist (template)
#25	Evidence summary table (example)
#26	QAS report drafting checklist (template)
#27	Introductory statement to the CPC (example)
#28	Annotated evaluation report (example)
#29	Lessons learned summary document (example)
#30	OIOS/IED documents request (example)
#31	Project Summary – Evaluation of the UN Department of Safety and Security (example)
#31a	Project Summary – Evaluation of Protection of Civilians in UN Peacekeeping Operations (example)
#33	Interview Guide Subprogramme Stakeholder (example)
#34	Focus Group Guide (example)
#35	Letter to Mission Head when conducting Member State surveys (example)
#36	Letter to Executive Committee when conducting a survey in English (example)
#37	Letter to Executive Committee when conducting a survey in French (example)
#38	Letter to Governing Council members when conducting a survey (example)
#39	Survey questionnaire - DESA staff (example)
#40	Survey questionnaire - Executive Committee (example)
#41	Survey questionnaire - Governing Council (example)
#42	Survey questionnaire - UNEP staff (example)

Relevant Websites - If a given hyperlink does not take you to the desired website, copy and paste the link into your web browser.

Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP)	www.alnap.org
Australian Bureau of Statistics	http://www.abs.gov.au/AUSSTATS/abs@.nsf/2f762f95845417ae-ca25706c00834efa/b75fd55ff9f9481bca2571fe007d69e3!OpenDocument
BetterEvaluation	http://www.betterevaluation.org
Engendering Policy through Evaluation	http://www.feministevaluation.org/
Gender and Evaluation	http://gendereval.ning.com/
Integrated Monitoring and Documentation Information System (IMDIS) data	http://imdis.un.org/
United Nations Joint Inspection Unit (JIU)	https://www.unjiu.org/
Medical Education Online	http://www.med-ed-online.org/rating/reliability.html
Mixed Methods International Research Association (MMIRA)	http://mmira.wildapricot.org/
My M&E	http://www.mymande.org/
National Center for Biotechnology Information	http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3402032/
OIOS Annual Reports	http://www.un.org/Depts/oios/pages/annual_reports.html
OIOS Programme Manager Reports	http://www.un.org/Depts/oios/pages/pm_reports_july12.html
OIOS-IED evaluation reports	http://www.un.org/Depts/oios/pages/other_oios_reports.html
Other UN governing body official documents	http://research.un.org/en/docs/symbols
Random number generators	http://stattrek.com/statistics/random-number-generator.aspx

Research Rundowns	http://researchrundowns.wordpress.com/quantitative-methods/signifi-cance-testing/
Sample size calculators	http://www.surveysystem.com/sscalc.htm
S-G reports to UNGA and/or UN Security Council	http://research.un.org/en/docs/symbols_
Social Research Methods	http://www.socialresearchmethods.net/kb/sampnon.php http://www.socialresearchmethods.net/kb/relandval.php
Sophia Learning	http://www.sophia.org/tutorials/simple-random-sampling-srs
Special Committee on Peace-keeping Operations (C34)	http://www.un.org/en/peacekeeping/ctte/CTTEE.htm
Stat Trek	http://stattrek.com/sampling/simple-random-sampling.aspx
Tablet-based note-taking	http://notesplusapp.com/ http://www.phatware.com/
Time zone difference calculator	http://www.timeanddate.com/
United Nations Abbreviations and Acronyms	http://dd.dgacm.org/editorialmanual/ed-guidelines/style/abbreviations.htm
United Nations Board of Auditors (BOA)	http://www.un.org/en/auditors/board/
UNGA Committee for Programme and Coordination (CPC)	http://www.un.org/en/ga/cpc/
United Nations Conference on Trade and Development (UNCTAD)	http://unctad.org/en/Pages/Home.aspx
United Nations Department of Economic and Social Affairs (DESA)	http://www.un.org/en/development/desa/index.html
United Nations Department of Field Support (DFS)	http://www.un.org/en/peacekeeping/about/dfs/

United Nations Department for General Assembly and Conference Management (DGACM)	http://www.un.org/depts/DGACM/
United Nations Department of Management (DM)	http://www.un.org/en/hq/dm/
United Nations Department of Peacekeeping Operations (DPKO)	https://www.un.org/en/peacekeeping/about/dpko/
United Nations Department of Political Affairs (DPA)	http://www.un.org/wcm/content/site/undpa/
United Nations Environment Programme (UNEP)	http://www.unep.org/
United Nations Evaluation Group (UNEG)	http://www.uneval.org/
United Nations General Assembly - Fifth Committee	http://www.un.org/en/ga/fifth/
United Nations General Assembly and UN Security Council resolutions	http://research.un.org/en/docs/symbols
United Nations High Commissioner for Human Rights (OHCHR)	http://www.ohchr.org/EN/Pages/WelcomePage.aspx
United Nations High Commissioner for Refugees (UNHCR)	http://www.unhcr.ch/
United Nations Human Settlements Programme (UN-Habitat)	http://www.unhabitat.org/categories.asp?catid=9
United Nations Independent Audit Advisory Committee (IAAC)	http://www.un.org/ga/iaac/
United Nations Office for the Coordination of Humanitarian Affairs (OCHA)	http://www.unocha.org/
United Nations Office of Human Resources Management (OHRM)	http://www.un.org/staffdevelopment/viewpage.asp
United Nations Office of Legal Affairs (OLA)	http://legal.un.org/ola/
United Nations Office on Drugs and Crime (UNODC)	https://www.unodc.org/
United Nations Office of Internal Oversight Services (OIOS)	http://www.un.org/Depts/oios/
United Nations OIOS USG	http://www.un.org/Depts/oios/pages/usg.html
United Nations organizational structure	http://www.un.org/en/aboutun/structure/org_chart.shtml
United Nations peacekeeping operations (PKOs)	http://www.un.org/en/peacekeeping/operations/current.shtml
United Nations Secretariat departments and programmes	http://www.un.org/en/mainbodies/secretariat/
United Nations special political missions (SPMs)	http://www.un.org/wcm/content/site/undpa/main/about/field_operations
United Nations terminology	http://untermportal.un.org/portal/welcome
WikiHow	http://www.wikihow.com/Assess-Statistical-Significance

External References and Further Reading

Links to the following documents can be found on the OIOS website at
http://www.un.org/Depts/oios/pages/iedmanualresources.html

Brinkman and Steiner, 2014a	Brinkmann, Svend, and Kvale Steiner. 2014. Interviews - Learning the Craft of Qualitative Research Interviewing (no electronic document)
Chow. 1996a	Chow, Siu L., 1996. Statistical Significance: Rationale, Validity and Utility, Volume 1 of series Introducing Statistical Methods
Cohen. 2011a	Cohen, N., & Arieli, T. 2011. Field research in conflict environments: Methodological challenges and snowball sampling. Journal of Peace Research, 48(4), 423-435 (no electronic document)
Cook et al. 1979a	Cook, T.D., and Campbell, D.T. 1979. Quasi-Experimentation: Design & Analysis Issues for Field Settings (no electronic document)
Daniel. 2011a	Daniel, J.N. 2011. Sampling Essentials: Practical Guidelines for Making Sampling Choices
Dillman et al. 2008a	Dillman, Don A., J.D. Smyth, and L. Melani Christian. 2008. Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method (no electronic document)
Ford et al. 2009a	Ford, N., Mills, E., Zachariah, R., et ux. 2009. Ethics of conducting research in conflict settings, Conflict and Health, 3:7 (no electronic document)
Gonick et al. 1993a	Gonick, L., and Smith, W. 1993. The Cartoon Guide to Statistics. (no electronic document)
Hair et al. 2009a	Hair, J. F., Tatham, R.L., Anderson, R.E., and Black, W. 2009. Multivariate Data Analysis (7th Edition). (Section IV, Chapter 8: Cluster Analysis) (no electronic document)
Hancock et al. 2011a	Hancock, D., and Algozzine, B. 2011. Doing Case Study Research: A Practical Guide for Beginning Researchers, Second Edition (no electronic document)
Henry. 1990a	Henry, G. 1990. Practical Sampling (no electronic document)
Kline. 2004a	Kline, Rex. 2004. Beyond Significance Testing: Reforming Data Analysis Methods in Behavioral Research Washington, DC: American Psychological Association
Levy et al. 2013a	Levy, P.S., and Lemeshow, S. 2013. Sampling of Populations: Methods and Applications
OECD DAC. 2009a	OECD DAC. 2009. Glossary of Key Terms in Evaluation and Results Based Management
Quinn Patton, Michael. 2008a	Quinn Patton, Michael. 2008. Utilization-focused evaluation (no electronic document)
Rockefeller Foundation. 2012a	Bamberger, Michael. 2012. Introduction to mixed methods in impact evaluation
Shannon et al. 2012a	Shannon et al. 2012. Emerging Themes in Epidemiology - Choosing a survey sample when data on the population are limited: a method using Global Positioning Systems and aerial and satellite photographs
Stake. 1995a	Stake, R.E., 1995. The Art of Case Study Research (no electronic document)
Trinidad. 2008a	Government of Trinidad and Tobago. 2008. Monitoring and Evaluation Division, Ministry of Social Development
Tourangeau. 2014a	Tourangeau, R., Edwards, B., Johnson, T., et ux. (Eds.), 2014. Hard-to-Survey Populations (no electronic document)
Tufte. 1990a	Tufte, Edward. 1990. Envisioning Information (no electronic document)
Tufte. 1997a	Tufte, Edward. 1997. Visual Explanations: Images and Quantities, Evidence and Narrative (no electronic document)
Tufte. 2001a	Tufte, Edward. 2001. The Visual Display of Quantitative Information (no electronic document)
Tufte. 2006a	Tufte, Edward. 2006. Beautiful Evidence (no electronic document)
UN Women. 2010a	UN Women. 2010. Manager's Guide to Gender Equality & Human Rights Responsive Evaluation

UNEG. 2005a	United Nations Evaluation Group. 2005. Norms for Evaluation in the UN System	UNEG. 2013b	United Nations Evaluation Group. 2013. Impact Evaluation in UN Agency Evaluation Systems: Guidance on Selection, Planning and Management
UNEG. 2005b	United Nations Evaluation Group. 2005. Standards for Evaluation in the UN System	UNEG. 2013c	United Nations Evaluation Group. 2013. UNEG Handbook for Conducting Evaluations of Normative Work in the UN System
UNEG. 2008a	United Nations Evaluation Group. 2008. Core Competencies for Evaluators of the UN System		
UNEG. 2008b	United Nations Evaluation Group. 2008. UNEG Core Competencies for Heads of Evaluation Offices in the United Nations	UNICEF. 2011a	UNICEF. 2011. How to design and manage Equity-focused evaluations
		United Nations. 1982a	A/37/38 of 9 June 1982 entitled Report of the Committee for Programme and Coordination on the work of its twenty-second session
UNEG. 2008c	United Nations Evaluation Group. 2008. UNEG Job Description for Evaluators in the UN System - Associate Evaluation Officer P1-P2		
UNEG. 2008d	United Nations Evaluation Group. 2008. UNEG Job Description for Evaluators in the UN System - Evaluation Officer, P3	United Nations. 1982b	A/RES/37/234 of 21 December 1982 entitled programme planning
		United Nations. 1984a	United Nations. 1984. A Guide to Writing for the United Nations
UNEG. 2008e	United Nations Evaluation Group. 2008. UNEG Job Description for Evaluators in the UN System - Intermediate Evaluation Officer, P4	United Nations. 1994a	A/RES/48/218 B of 29 July 1994 entitled Review of the efficiency of the administrative and financial functioning of the United Nations
UNEG. 2008f	United Nations Evaluation Group. 2008. UNEG Job Description for Evaluators in the UN System - Senior Evaluation Officer, P5	United Nations. 1994b	ST/AI/397 of 7 September 1994 regarding Reporting of inappropriate use of United Nations resources and proposals for improvement of programme delivery
UNEG. 2010a	United Nations Evaluation Group. 2010. UNEG Quality Checklist for Evaluation Terms of Reference and Inception Papers	United Nations. 1994c	ST/SGB/273 of 7 September 1994 entitled Establishment of the Office of Internal Oversight Services
UNEG. 2010b	United Nations Evaluation Group. 2010. UNEG Quality Checklist for Evaluation Reports	United Nations. 1995a	ST/AI/401 of 18 January 1995 entitled Personnel arrangement for the Office of Internal Oversight Services
UNEG. 2010c	United Nations Evaluation Group. 2010. Good Practice Guidelines for Follow up to Evaluations	United Nations. 1997a	ST/SGB/1997/5 of 12 September 1997 entitled Organization of the Secretariat of the United Nations
UNEG. 2011a	United Nations Evaluation Group. 2011. Integrating Human Rights and Gender Equality in Evaluation - Towards UNEG Guidance	United Nations. 1998a	ST/SGB/1998/2 of 12 February 1998 entitled Organization of the Office of Internal Oversight Services
		United Nations. 1999a	A/RES/54/244 of 23 December 1999 entitled Review of the implementation of General Assembly resolution 48/218 B
UNEG. 2013a	United Nations Evaluation Group. 2013. UNEG Strategy 2014-2019		

United Nations. 2000a   ST/SGB/2000/8 of 19 April 2000 entitled UN Regulations and Rules Governing Programme Planning, the Programme Aspects of the Budget, the Monitoring of Implementation and the Methods of Evaluation (PPBME)

United Nations. 2002a   ST/SGB/2002/7 of 16 May 2002 entitled Organization of the Office of Internal Oversight Services

United Nations. 2004a   A/RES/59/272 of 23 December 2004 entitled Review of the implementation of General Assembly resolutions 48/218 B and 54/244

United Nations. 2007a   A/RES/61/245 of 7 March 2007 entitled Comprehensive review of governance and oversight within the United Nations and its funds, programmes and specialized agencies

United Nations. 2007b   A/RES/61/275 of 31 August 2007 entitled Terms of reference for the Independent Audit Advisory Committee and strengthening the Office of Internal Oversight Services

United Nations. 2010a   A/64/263 of 29 March 2010 entitled Review of the implementation of General Assembly resolutions 48/218 B, 54/244 and 59/272

United Nations. 2013a   A/RES/68/21 of 11 December 2013 entitled Report on the activities of the Office of Internal Oversight Services

USAID. 2013a   USAID. 2013. Technical Note – Conducting Mixed Methods Evaluations

Urdan. 2010a   Urdan, T. 2010. Statistics in Plain English, Third Edition (no electronic document)

World Bank. 2009a   Morra Imas, Linda G., and Ray C. Rist. 2009. Road to Results. Designing and Conducting Effective Development Evaluations

World Bank. 2011a   World Bank Independent Evaluation Group (IEG). 2011. Writing terms of reference for an evaluation: A how-to guide

Yin. 2011a   Yin, R.K. 2011. Applications of Case Study Research (Third Edition) (no electronic document)

Yin. 2013a   Yin, R.K. 2013. Case Study Research: Design and Methods (Fifth Edition) (no electronic document)