

A REVIEW OF

COUNTRY PROGRAMME

EVALUATIONS

IN SELECTED UN ORGANIZATIONS AND

INTERNATIONAL BANKS

Commissioned by

UNICEF
Division of Evaluation, Policy and Planning

Olivier Cossée
Monitoring & Evaluation Consultant

15 March 2001

Table of Content

1. Background	1
2. Rationale.....	2
3. Experience of Selected Organizations	2
3.1. UNICEF.....	2
3.2. UNDP	3
3.3. FAO.....	4
3.4. WFP.....	5
3.5. World Bank	5
3.6. Inter-American Development Bank (IADB)	7
4. Current Trends and Issues	11
5. Findings and recommendations	12
5.1. The Timing Dilemma	12
5.2. Finding a Focus	13
5.3. How to maximize CPE usefulness for the next CP?.....	14
5.4. External vs. Internal Evaluators.....	15
5.5. Interviewing Beneficiaries in CPEs.....	15
5.6. Looking for Impact.....	16
5.7. Blending Review and Evaluation	17
5.8. How to Use Standard TORs?.....	18
5.9. Scheduling the Review	18

1. Background

Since the 1980's, many multilateral aid organizations have had to reform their ways of operation, often under the pressing request of their respective governing bodies. As it became apparent that UN interventions at the country level were often scattered and lacking a coherent frame of reference, one idea that gained currency in the mid-1980's was country programming. The idea is basically to draft a frame or reference – or master plan of operation, or country programme – to ensure that aid provided by a given organization to a given country is relevant to its needs and opportunities; and to provide structure, strategic vision and coherence to what was so far a mere portfolio (collection) of independent projects. In many organizations country programming was linked to increased decentralization of management authority to their country or regional offices.

Piloted by a few organizations including UNICEF, country programming soon spread to become a general practice, almost an evidence. Obviously the terminology varies from one organization to another: country programmes, country papers, country assistance strategy, master plan of operation are but a few names given to what is basically a document describing in a summary fashion the main sectors of intervention, strategies, objectives and financial resources available through a given aid organization, in a given country and for a given period. Prepared at the country level through negotiations with national authorities, the country programme is customarily approved by the concerned government and the executive board of the organization.

Country programme evaluations are a direct consequence of country programmes. Half a decade after pushing successfully for country programming, Executive Boards asked for evaluations that would follow the same structure and account for CP resources. Again, multilateral and bilateral aid agencies use different terms and formats for the purpose of accounting for country programmes. Indeed quite a variety of formats and methodologies are being tested and refined.

In March 1999, the OECD Development Assistance Committee organized a workshop on country programme evaluations in Vienna, to take stock of methodological advances and discuss the experience among bilaterals and IFIs. The workshop highlighted the growing popularity of CPEs as well as a wide variety of methods and formats. A number of critical issues discussed in the workshop have been included in the questionnaires for the present review.

UNICEF was one of the first UN organizations to implement Country Programmes and Country Programme Evaluations. The first pilots, undertaken in the mid-1990's, did not lead to the broad use of CPEs as conceived by the Evaluation and Research Office. A lighter procedure, the end-of-cycle review, took over the CP accountability function. Recently the Division of Evaluation, Policy and Planning (EPP) has embarked on a revision of its modus operandi, and the CPE is a formula that might be revisited.

Acknowledging that any new attempt need to take into account the lessons of the first UNICEF pilots as well as the experience of similar international aid agencies, EPP commissioned a small review to extract preliminary lessons and best practices from the experience of selected UN agencies and International Finance Organizations.

The review focused on the CPE experience of UNICEF, UNDP, FAO, WFP, IADB and the WB. The consultant reviewed CPE-related documentation and interviewed staff in central evaluation offices in New York, Washington and Rome. The proceedings of the DAC Vienna Workshop¹ were also a source of information and inspiration for the next section.

Before describing the experience of selected organizations, it is important to explore, as summarily as possible, the purpose of CPEs.

¹ OECD-DAC: Evaluating Country Programmes, Vienna Workshop, 1999.

2. Rationale

As any evaluation exercise, a CPE aims at three objectives:

- 1) *Account* to the Board (and donors) on results achieved under a Country Programme approved by the same Board a few years before;
- 2) *Improve* the relevance, efficacy and efficiency of subsequent CPs through a series of recommendations feeding into CP preparation; and
- 3) *Extract* general lessons on aid effectiveness, factors influencing performance, the implementation of corporate strategies, etc.

In practice, the first objective (accountability) is the dominant reason why central evaluation offices are pushing for CPEs, in response to a very unanimous demand from their respective Executive Board. The second and more forward-looking objective (contribute to the next CP formulation) is given increasing importance. It is of course a key objective if CPEs recommendations are to be implemented at the country level.

The third objective is probably the most ambitious. In this perspective CPEs are viewed as a unit of data gathering and analysis. None of the reviewed organization has yet reached this stage. The lessons one can extract from a CPE are often country-specific. It looks like CPEs are too general in scope and too specific in location to produce many general lessons. Thematic evaluations, with their ability to compare a single strategy or approach applied in different countries, are probably the best-suited tools to address the third evaluation objective.

A number of rationales have been proposed for country-level evaluations of aid effectiveness. It has been said that within a given country, “socio-economic and natural parameters reach a certain degree of homogeneity and permanence” and that therefore “the country is in most cases the most logical unit of aid management and account”.²

The fact of the matter is that the country is a key level in aid management. The government is a key player in any multilateral cooperation programme. Within aid agencies, decentralization has invested country representations with increased resource management authority. And if decisions are made primarily at the country level, it makes a lot of sense to evaluate at the same level the consequences of these decisions. Socio-economic conditions vary in any country, but as long as the country is an important unit of aid management, it will remain an important unit of aid evaluation.

3. Experience of Selected Organizations

The extent to which selected agencies have experimented with CPEs depends directly on how long they have been using country programming. Country programming started late in some agencies (WFP, IADB) and those are still at an early stage of their involvement with CPEs. Others (UNICEF, UNDP) adopted CP sooner and were already experimenting with CPEs in the early 1990's. However they often encountered problems related to the scope of such evaluations. Early CPEs tried to embrace too much, both from a topical point of view (long TORs) and in terms of “breadth” or comprehensiveness (covering most projects and programmes under a given CP). Most agencies are now trying to develop a lighter format with more focused TORs and only a sample of projects coming under review. The experimentation phase is far from over, but the corpus of experience is much larger than a few years ago.

3.1. UNICEF

Soon after the adoption of Country Programming, UNICEF started experimenting with ways of assessing CP performance. Annual and end-of-cycle reviews were established to take stock of implementation

² OECD-DAC, Op. Cit.

progresses/challenges and plan for the period ahead. A very operation-oriented tool, implementation reviews are mandatory and – at least in principle – managed by the recipient government. In practice the regional management team plays a large role in CP reviews.

In 1992, a Multi-Donor Review of UNICEF stated that end-of-term reviews were insufficient and called for a more evaluative approach. CP Evaluations were developed in the early 90's as a tool distinct from reviews, to build up CP accountability in more evaluative areas such as CP relevance, impact, coverage, cost-effectiveness and sustainability of results. Standard TORs were developed by ERO and 7 trial CPEs were conducted during 1993 and 1994 (Argentina, Brazil, China, Egypt, Malawi, Philippines, Thailand). Some of these were stand-alone exercises, others were performed in conjunction with annual or end-of-cycle reviews.

In 1995, the experience gathered in trial CPEs was analyzed by ERO in a series of documents³ in which the following issues are identified:

- The original intent was to compare CP goals and achievements, but this was made difficult by limited available database on results and imprecise and/or changing goals; consultants thus tended to focus on softer, more qualitative issues such as CP continued relevance;
- The process of drafting specific TORs has led to a broad scope (long lists of evaluation issues) and a subsequent lack of focus in CPEs.
- The tested duration of one or two man.month was not sufficient to cover activities and results at the project level, did not provide opportunity for ground-truthing reports and interviews with programme staff, nor for field visits and primary data collection;
- Consultants often had to narrow down the focus themselves, either on their own discretion or following discussions with the CO. The choices made were not always pertinent and this has resulted in uneven, not comparable reports;
- The relationship with end-of-cycle reviews was unclear, there was perceived overlap between the two instruments; unlike reviews, CPEs lacked a direct connection with the planning process;

To address these issues, the 1995 ERO review recommended:

- greater clarity vis-à-vis end-of-cycle reviews
- greater participation of COs in CPEs
- to explore possibilities to interview beneficiaries
- to time the CPE carefully, taking into account planning needs (tension between accountability and planning objectives)
- Improve criteria for choice of countries; include some emergency countries
- Vast scope means that management and evaluators have to make pertinent choices.

A few more CPEs were conducted in 1996 (Chad, Haiti, Chile) and 1997 (Senegal), but the 1995 review did not really lead to generalization of CPEs in UNICEF. Some “fuzzy” issues could not be resolved and lead to the demise of the CPE tool in UNICEF. Chief among them is the central issue of the relationship with end-of-cycle reviews, which continued to this date and seem to have taken on some evaluative dimension (e.g. Togo, 1997). The current feeling in the Division for Evaluation, Policy and Planning is that an improved end-of-cycle review would serve the purpose of CPEs better than a stand-alone instrument.

3.2. UNDP

UNDP briefly experimented with CPEs in the mid-90s (Sri Lanka 1994, Niger 1995, Uganda 1996). The idea was to develop a standard evaluation procedure for country programmes. The attempt was short-lived however. CPEs had a number of disadvantages compared to Mid-Term Reviews. The MTR was a

³ CPE Briefing Note, ERO March 95; Country Programme Evaluations Summary Review, Sam Bickel, Oct. 95; CPE Guidelines, ERO, Dec. 95.

mandatory management step strongly supported by the Board and scheduled soon enough to contribute to the planning process for the coming CP. CPEs on the other hand were non-mandatory exercises coming at the end of the cycle and therefore could not contribute to the planning process for the next CP.

The original CPE idea has now been abandoned in UNDP, in favor of a lighter exercise: the Country Cooperation Framework (CCF) review, which replaced the MTR. The responsibility for country reviews has recently been given to the Evaluation Office after being under an operational unit for some years. The EO was selected because it is leading the Result Based Management (RBM) exercise. Since 1999, RBM is viewed as essential to focus the organization on demonstrable results. CP reviews could help validate from an independent evaluator's perspective the self-assessment work done under RBM procedures. CCF reviews are currently being re-designed by the EO in order to be more outcome-oriented.

CCF reviews are strongly supported by the Board, which requires a minimum of 10 per year. About 20 CCF reviews were presented to the Board in 1999 and 70 in 2000.

Long standard TORs are available. The Evaluation Office wants to promote a modular approach to TOR definition. Different sizes of country reviews are considered: a "light" review covering part of the TORs and managed by the Country Office, and a full-scope review, covering most of the standard TORs, piloted by EO and presented to the Board. Another idea along the same lines was the recent cluster review of 8 small CPs in Eastern Europe.

CCF reviews are viewed as a way to report not only on projects, but also on the entire UNDP "presence" (coordination role, advocacy, policy dialogue, political initiatives, etc.). However, they have not been very successful on advocacy and soft interventions in general. The TORs also cover outcomes and impact. In practice country reviews do not do justice to the outcome level, let alone impact. It is quite a light exercise: UNDP uses one or two independent consultants or (quite often) ex-staff members for about one month each to conduct a review. There are no beneficiary interviews and little fieldwork in general. The quality of the reports varies. Reliance on ex-staff members has been mentioned as a cause for the low quality of some reports.

However, the EO is thinking of using PRA, focus group interviews and beneficiary surveys, as in the CLIA methodology tested in 2000 in Burkina Faso and Malawi. The main idea behind the methodology was to apply participatory evaluation techniques to impact assessment. The evaluators would collect impact testimonies from all stakeholders of a few sampled projects, compare them with already documented impact-related information (from evaluation reports, monitoring systems, research), and draw some conclusions in the form of documented impact assumptions.

One of the idea successfully tested in CLIA to "cut down the chase" for impact was to start with impact itself, i.e. experienced change at the country, institutional or community levels, and to make one's way back to the CP contribution if there is any. This sort of retro-analysis is faster than the traditional way of doing impact assessment, which is the other way around: list projects and make your way up to results, impact and sustainability for each project. Retro-analysis is very "macro" though, and does not detect micro-level and preliminary impacts. It works better with visible, "hardware" projects (service delivery, infrastructures, institution building, etc.) than with more subtle, "software" activities (policy advisory, rights and cultural issues, etc.). Retro-analysis can obviously be combined with classic "forward-analysis" for improved results.

Another, less successful idea was that CLIA impact assumptions would be tested subsequently through rigorous data collection protocols and surveys. This "research" phase, supposed to validate the results of the first rapid assessment phase with fresh data, requires more time and resources than normally required for a CP evaluation.

3.3. FAO

FAO has faced a funding decline in recent years that translated into important downsizing of country-level operations and the adoption of a new focus on normative work and inter-governmental coordination on key issues and international conventions. There is not much to say about the agency involvement in CPEs, except for a short period when FAO was participating in UNDP CP Evaluations. The experience is limited to the end of 1980's – early 1990's: Nepal (1988), Mozambique (early 1990's) and a few other countries. Formal CPEs have not been performed since, partly because the relationship with UNDP loosened as a result of increased use of national execution, and partly because the funding crisis means there is not much programming going on at the country level anyway.

The agency has been increasingly active in emergency situations (Iraq, Kosovo) and country-wide evaluations have been conducted in this particular context. One could argue that since there is no time to make a CP in an emergency, it is difficult to talk of country programme evaluations in these cases. There is no comparison with pre-set goals at the country level, so the exercise resembles more a country case study or a portfolio assessment than a CPE. However, the distinction is weaker in practice than in theory because CP goals are often too general in wording to allow for a straightforward comparison between goals and achievements. And even an emergency operation can be planned and have clear objectives with or without a CP. Emergency country evaluations are very close to CPEs. The reason why FAO is doing the former and not the latter is simply that emergency countries are currently the only places where FAO can mobilize a critical mass of resources for country programming.

CPEs make perhaps less sense to specialized agency because each of their CP remains fairly sectoral by nature and therefore less interesting to analyse country-wide and cross-sectorally. FAO is more interested in thematic evaluations (called “technical programmes evaluations”, such as the one on the Special Programme for Food Security to be launched in 2001) through which one can more easily learn lessons in technical and specialized fields.

3.4. WFP

CPEs are rather new to WFP, as is the principle of country programming itself, introduced in 1994. The evaluation office is currently developing a new vision for its work and CPEs are on the rise, but the office acknowledges that quality and methodology should be improved. One CPE was conducted in 1999, 8 in 2000. Half of all evaluations in 2001 should be CPEs. Ultimately, it is envisaged to transfer CPEs responsibility to Country Offices.

CPEs respond to a Board requirement. They are normally presented to the Board at the same meeting than the new CP strategic outline. The focus is not so much on accountability, i.e. on what has happened as a result of the CP. Rather, the standard TORs are concentrating on design issues such as CP relevance and general processes, conformity to the new mandate of the organization, fit between CP and country situation, and recommendations for new CP positioning. Outcome & impact are deemed too difficult to cover. The project level is explicitly excluded from the standard TORs, which deal almost exclusively with the CP level.

The country focus and resource allocation process in WFP is quite flexible and depends on changing local food security situations and unpredictable crisis. As in FAO, there seems to be demand for the review of emergency operations on a country basis (Kosovo & North Korea in 2000). Again and despite the absence of a CP, these emergency country evaluations are quite close to CPEs and reflect the value-added of evaluating agency operations at the country level in emergency countries as well.

3.5. World Bank

The CAS (Country Assistance Strategy, equivalent to a CP) has become a key programming step in the World Bank. The document is obviously submitted to, and approved by the Board. In this context, the CA Evaluation (CAE) has become since 1995 the main accountability mechanism. “When the CAS came, the Board did not find project evaluations so relevant anymore”. The CAE is typically presented to the Board

one month prior to the new CAS. OED is responsible for preparing CA evaluation reports and presenting them to the Board.

The Bank has finalized a dozen CAEs (Argentina, Bulgaria, Ghana, Ivory Cost, Khazakstan, Morocco, Mozambique, Philippines, Poland, Tanzania, Zambia...). The current rhythm is about seven per annum (ongoing: Peru, Lesotho...). Ultimately, the aim is to review a CAS every one and ½ year.

The World Bank experimented with CAEs large and small. The small ones were known as CA Notes (CANs) and the large ones as CA Reviews (CARs), but the current terminology is simply CAEs of various sizes. What seems to vary between large and small ones are the length of contextual material and the depth of the cause-to-effect analysis (more information on outcomes and impact in large CAEs). A first CAE is also expected to be larger than subsequent updates. The cost varies from \$50,000 to \$500,000.

CAEs are usually conducted at mid-term. The period covered does not depend on the planning cycle. It is typically longer (about 5 to 10 years) than a CAS. The starting point for the review period is usually a key turning point in the country history.

The topical span is large: issues covered include the macro-economic situation, Bank performance in lending and non-lending work (a.k.a. “analytical and advisory”), government performance, stakeholders’ views on Bank support, outcomes and impact, and other donors’ work.

There is quite a refined methodology. Outcomes and impacts are routinely described in qualitative terms by sector of intervention. To assess results and impact, the Bank uses 3 types of comparators (CAS objectives, macro-economic forecasts and neighboring countries development trajectory), a project rating grid (ratings by evaluators) and structured stakeholder interviews. Interviewed stakeholders include other donors, Ministry of Finance (counterpart), program staff, beneficiaries, and academia. Program evaluations (“audits”) feed into the process.

The Bank pioneered project rating, but had a word of caution during the 1999 Vienna Workshop⁴: evaluators’ subjectivity cannot be completely overcome. The best one can achieve is comparability within a given country. Comparability is much more difficult to achieve between countries.

Government involvement in the review is generally low, limited to getting a copy of the draft report for comments. Tanzania and Bulgaria were mentioned as examples of increased recipient country participation, in an attempt to achieve higher national ownership of evaluation findings.

Amongst the reviewed documentation, the WB CAE are probably the best CPEs. The reports include a vivid context description and historical section. The analysis focuses on implementation progress and outcomes, and is quite thorough at least on the former. Wording can be harsh at times. Annexes include projects summary sheets for all reviewed projects⁵ or at least a listing of projects reviewed⁶, as well as a formal “management response” to the evaluation (position of program managers against each recommendations) and a report from the Committee on Development Effectiveness giving additional feedback on how the review was received and used in the Bank.

A strong emphasis is placed on contextual information in the CAE process and documents, a useful characteristic when it comes to evaluating CAS relevance and macro-economic results over time. Former managers in the Bank and the government are traced and interviewed as part of a deliberate attempt to “get the context right”. A last lesson from WB experience is that systematization of the process and a consistent methodology are key to getting results accepted at the country level (“buy-in”).

⁴ OECD-DAC: Evaluating Country Programmes, Vienna Workshop, 1999, p. 23.

⁵ Mozambique CAR.

⁶ Philippines CAR.

3.6. Inter-American Development Bank (IADB)

Country programming itself is relatively new to IADB. Country Papers (CPs, equivalent to a Country Programme) were said to have encountered some resistance and were sometimes perceived by countries as straight-jackets and/or a vehicle for the Board's agenda (e.g. poverty or governance projects). Here again, CPEs meet a direct Board request to monitor CP implementation. Promoted by the Evaluation Office as early as 1996, the CPE has since become a mandatory step in the programming process and remains to this date under the responsibility of the Evaluation Office.

In 1995, the IADB started a series of CP "evaluability exercises" in order to assess whether the CPs presented to the Board had goals and strategies described "in such a way as to enable evaluation" (precisely enough, with indicators, baseline, etc.). The result was that "in most cases, the CPs were unclear as to how the achievements of the goals and its objectives were to be measured at the end of the cycle".

The Evaluation Office undertook two pilot CPEs in 1997 (Equator and Dominican Republic). There were not problem-free. The fact that CPs lacked clear measurable goals was a constraint. The evaluations were also faced with some resistance from operational staff who did not really "buy-in". The next generation of CPEs (Peru, Trinidad & Tobago in 2000, Honduras, Nicaragua, Mexico, Guyana, Bolivia in the pipeline, goal of three or four CPEs per year) aims at a greater involvement of operational divisions in the process, starting with the choice of a CPE focus.

Typically presented to the Board three or four months before the new CP presentation, the CPE covers a long period, about five to eight years (two cycles). There are no standard TORs, since flexibility in setting the evaluation focus is perceived as key to its success. CPEs usually cover the economic growth & poverty context, relevance of CP, quality of diagnosis, positioning, implementation & key projects, and CP outcomes. Each evaluation costs between \$100,000 and \$200,000.

As far as impact is concerned, the typical evaluation just uses already documented data. Project evaluations feed into the CPE. The methodology also includes focus group interviews of main stakeholders, beneficiary interviews for a sample of important projects (they wish they could do more), and interview of past Bank & country management teams.

According to staff in the IADB, the focus on a few relevant issues and projects is key to success. Humility is necessary. The risk of writing encyclopedias should be emphasized and controlled through the carefully selection of a focus based on the CP itself and responding to operational staff information needs.

Table 1: Summary Experience of Agencies

Agency	Context	Rationale	Focus	Methodology	Experience so far	Current Trends / Lessons
UNICEF	<ul style="list-style-type: none"> - Master Plan of Operation (MPO) encapsulate CP goals - Decentralization perceived as one of UNICEF strengths - Important data collection at the country level (IMEP) 	<p>CP Evaluations and reviews are primarily a Board requirement.</p>	<ul style="list-style-type: none"> - Mainly focusing on positioning and implementation (processes). - Outcomes covered in best reviews. Impact not really covered so far. - Should cover the whole strategy mix: service delivery, capacity building and empowerment 	<ul style="list-style-type: none"> - A fairly light exercise: usually 2/3 man.month (consultant) - Standard TORs expanded in 95, now quite long but difficult to follow. 	<p>94-95: a dozen CPEs (Argentina, Brazil, China, Chile, Haiti, Ivory Coast, Malawi, Niger, Peru, Philippines, Senegal, Thailand, Togo, Uganda...) plus an unknown number of end-of-cycle reviews</p> <ul style="list-style-type: none"> - The particularity of CPEs as opposed to end-of-cycle reviews has never been clearly established and remained permeable. 	<ul style="list-style-type: none"> - How to elicit CO input has been an issue. - Feeling that a good end-of-cycle review should serve the purpose.
UNDP	<ul style="list-style-type: none"> - Since 1999, Result Based Management (RBM) is viewed as essential to focus the organization and stop funding decline. - The Country Cooperation Framework is subject to a mandatory final review. - CCF reviews now managed by Evaluation Office, as a check to RBM self assessment. 	<ul style="list-style-type: none"> - Requested by the Board, not much interested in project. Evaluations since the advent of the CP and CCF. - Take stock of new context. - Contribute to next CCF (especially positioning) - Could help validate RBM self-assessment. 	<p>Quite broad, with emphasis on relevance / positioning issues, projects & programmes implementation but also “soft interventions” (coordination & advisory roles), and outcomes whenever possible.</p>	<ul style="list-style-type: none"> - Timing: one year before end of CCF, to feed into next CCF formulation. - External missions raise the issue of CO and Government participation and ownership. - Outcomes not much covered, let alone impact, but format is being re-designed as more outcome-oriented. - Lang standard TORs. Different sizes are considered: a “light” review covering part of the TORs and done by CO, and a full-scope review, done by EO and presented to the Board 	<ul style="list-style-type: none"> - CPEs tried in the mid-90’s (Sri Lanka 94, Niger 95, Uganda 96) and abandoned because not linked to planning process. - The Board asked for Mid-Term Reviews (now CCF reviews). - About 20 CCF reviews in 1999 and 70 in 2000. - Participatory impact assessment experimented at the country level (CLIA) in Burkina Faso and Malawi (2000). 	<ul style="list-style-type: none"> - Reviews should be made more manageable and flexible. - Important to improve CCFs evaluability (with targets and clear intended outcomes) to improve reviews. - Thinking of using focus groups and beneficiary surveys in CCF reviews. - Idea of using reviews as building blocks for further analysis / comparisons.

Agency	Context	Rationale	Focus	Methodology	Experience so far	Current Trends / Lessons
FAO	Funding decline has meant a downsizing of country-level operations and a stronger focus on normative work and inter-governmental coordination on key global issues	Mainly done in conjunction with UNDP, participation in UNDP CP Evaluations	Relevance of CP/ projects, management processes and preliminary results.	Currently not performing CPEs	Limited to end of 80's – early 90's: e.g. Nepal (88) and Mozambique (early 90's) and a few others. CPEs have not been performed in recent years, because the relationship with UNDP loosened (NEX), and the funding crisis means little resources are available at the country level anyway (little country programming left).	Responsibility vested upon the COs, but demand seems almost non-existent at the moment, due to funding crisis. HQ evaluation office is focusing on thematic evaluations (called “technical programmes evaluations”) such as the one on the Special Prog. for Food Security to be launched in 2001. Also in demand: country-based evaluations of emergency operations (e.g. Kosovo in 2000)
WFP	Country programming introduced recently (1994). Size of WFP presence in countries depends on local food security situation / crisis	Basically a Board requirement. CPEs are presented to the Board at the same meeting than the new CP strategic outline. Accountability but also a way to push for the new mandate/focus.	CP relevance and general processes, relevance of projects, fit between CP and country situation, recommendations for new CP positioning. Outcome & impact deemed too difficult	Standard TORs focusing on CP design and positioning issues (relevance, quality, precision of objectives, formulation procedures, government involvement in CP preparation, etc.).	CPEs are rather new to WFP (because CPs are new). 1 in 1999, 8 in 2000. ½ of all evaluations in 2001 should be CPEs. Also: reviews of emergencies operations on a country basis (Kosovo & North Korea in 2000), which are not really CPEs because there's no CP involved.	HQ evaluation office is developing a new vision. CPEs are on the increase, quality & methodology should be improved. It is envisaged to transfer CPEs responsibility to COs, as a mid-term review.

Agency	Context	Rationale	Focus	Methodology	Experience so far	Current Trends / Lessons
World Bank (WB)	The CAS (Country Assistance Strategy, equivalent to a CP) has become a key programming step, submitted to the Board.	Accountability mechanism for the CAS (“When the CAS came, the Board did not find project evaluation so relevant anymore”). The CAE (Country Assistance Evaluation) is typically presented to the Board 1 month prior to the new CAS.	<ol style="list-style-type: none"> 1) Macro-economic situation 2) Bank performance (lending and non-lending work a.k.a. “analytical and advisory”) 3) Government performance 4) Stakeholders’ views on Bank support 5) Results of Bank support 6) Other donors’ work 	<ul style="list-style-type: none"> - Experimented with various sizes (=depth of review): CA Reviews (CARs) and CA Notes (CANs) - Period: start 1 year before new CAS is due to be presented at the Board, to submit CAR 1 month before. - Comparators: CAS objectives, macro-economic forecasts and other neighboring countries - Project ratings by evaluators - Questionnaires for stakeholders - Trace and interview former managers 	<ul style="list-style-type: none"> - Started in 1995, about 7 per annum now. Cost: \$50,000 to 500,000 - Finalized: Argentina, Bulgaria, Ghana, Ivory Cost, Khazakstan, Morocco, Mozambique, Philippines, Poland, Tanzania, Zambia... - Ongoing: Peru, Lesotho... - An effort to link up with other IFIS (IADB, ADB, IFC) in CARs 	<ul style="list-style-type: none"> - Very important to get the context right, and interviews of former managers are a useful tool for this. - Systematization of process and consistent methodology are key to acceptance of results and “buy-in”.
Inter-American Development Bank (IADB)	Country programming itself is relatively new to IADB. CPs sometime perceived as straight-jackets and/or a vehicle for the Board’s agenda (e.g. poverty or governance projects)	Basically a Board request to monitor CP implementation. CPEs have become a mandatory step in the programming process, presented to the Board 3 or 4 months before the new CP.	<ul style="list-style-type: none"> - Growth & poverty context and CP outcomes - Relevance, quality of diagnosis, positioning - Implementation & projects - Impact: only uses data already documented. - No standard TORs so far, flexibility in deciding focus is key 	<ul style="list-style-type: none"> - Project evaluations feed into the CPE - Interview of past Bank & country management teams - Focus group interviews of main stakeholders, beneficiary interviews for a sample of important projects - Cost: \$100,000 to 200,000 - Study long period, about 5 to 8 yrs (2 cycles) 	<ul style="list-style-type: none"> - First pilots in 97: Equator & Dominican Republic. Problem: no buy-in from operational staff. Also CPs lacked clear measurable goals. - 2000: Peru, Trinidad & Tobago - Pipeline: Honduras, Nicaragua, Mexico, Guyana, Bolivia (would like to do 3 or 4 per annum) 	<ul style="list-style-type: none"> - CPEs remain under EO, as a strong accountability requirement. - The focus on a few relevant issues and projects is key to success. Humility is necessary. Risk of writing encyclopedias. Carefully chose focus with operational staff, and based on CP. - Include more civic society agents in focus group interviews

4. Current Trends and Issues

The present section highlights a number of current trends on which most organizations agree, and some key issues that are still a matter of debate among them.

- **Almost all agencies like CPEs:** The only exception being FAO, which seems to favor thematic evaluations. But this may also reflect the size of their country-level resources, since even FAO is trying the format in their emergency programmes. Agencies that use country programmes to plan their assistance have some sort of CPE exercise at the core of their evaluation strategy.
- **Role of operational units and country offices:** has been weaker than it should have been. Evaluation offices are moving towards more demand orientation and greater involvement of operational units in order to build “buy-in” and evaluation usefulness. The issue of who should “own” or manage the CPE procedure in house (operational units or evaluation office) has been raised in many organizations. The best solution is perhaps shared ownership, but clearly, central evaluation offices have a strong role to play to ensure accountability, especially for large CPs.
- **Role of the government:** the agency – government working relationship seems to be the core of the matter as far as CPEs are concerned. As explained above, the fact that the government is a key partner and determinant of international aid provides the strongest rationale for CPEs. Issues such as national ownership, government performance and policy environment are central to the exercise. In terms of government ownership of, or participation in CPE themselves, the prevalent practice has been to consider the government more as the evaluated than the evaluator. But things are changing and the value of government adherence to the evaluation findings and recommendations is slowly being recognized.
- **Levels of analysis:** the CP level is the focus, but one can only document results by looking at projects and non-project activities in some details, so the project level should also be analyzed and reported in a summary fashion. A review of the strength and weaknesses of the portfolio would be useful to the formulation of the next CP. The issue of the linkages between CP and projects goals is still a bit fuzzy for some agencies.
- **Topical span:** relevance and positioning issues are the most extensively covered. Processes and short-term results come next. Agencies hold modest expectations as per the contribution CPEs could make to impact assessment. There was some discussion on this point in Vienna, but impact is often considered out of scope. The banks (WB & IADB) and UNDP compile in their CPEs whatever impact information has already been collected through project evaluations or studies.
- **Size of the exercise:** there is wide agreement that flexibility in defining the evaluation framework is key to keeping CPEs manageable. The exercise cannot be done cheaply. Four or five man.month seems to represent a minimum for any serious evaluative work at the country level.
- **Joint CPs:** multi-donor CPEs are not in the cards for now, although it would go in the same direction as the CCA and UNDAF. Most agencies are still experimenting and donor sensitivities come into play. Maybe a few pilots will be launched in the coming years.
- **Implicit and explicit CP:** the EU⁷ has made an interesting distinction between explicit CP (as written) and implicit CP (as implemented and/or as favored by the country), and used a comparison of spendings vs. plan as a quick way to highlight the discrepancies between the two. The concept can be useful in cases of systematic discrepancies between the planned and implemented CP, i.e. when one specific type of projects is consistently under-implemented while others are not.

⁷ Quoted from OECD/DAC, Op. Cit.

- **Ratings:** the WB has pioneered projects rating – a process of breaking down complex issues into elementary variables rated by evaluators – but acknowledges evaluators’ subjectivity and effect on ratings. Ratings can only come as complement to a detailed description. Comparability within country is adequate, but not between countries.⁸

5. Findings and recommendations

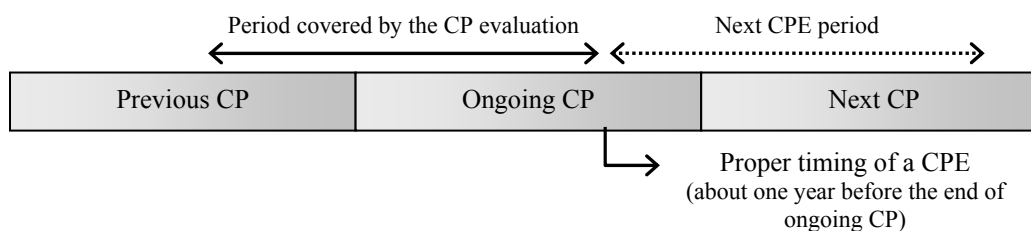
5.1. The Timing Dilemma

One of the recurrent problem faced by CPEs and the people who manage them is the issue of timing, where one need to strike a balance between the two main evaluation objectives of CPEs: accountability and contribution to planning. A CPE coming late in the planning cycle would be strong on accountability but would probably come too late to contribute significantly to the planning of the next CP, which has to be designed and approved before the current CP comes to an end. A CPE coming early in the cycle (e.g. at mid-term) would come early enough to help shape the future CP, but would find it difficult to document the results of the ongoing CP.

The only practical solution seems to be to **evaluate a longer period than the ongoing CP alone, i.e. include in the review the end of the previous CP and the transition between the previous and the current CP** (see Figure 1). This solution implies an evaluation period of five to eight years. It is often put to good use in IADB CPEs and WB CAEs. It has two main advantages:

- The surveyed period encompasses the transition period between the past and current CPs, and notably the design process for the current CP. This is only logical since a CPE should be looking at the relevance, quality of design, local ownership and participation of the CP. The evolution of sectoral/thematic focus, strategies and portfolio from the previous CP to the ongoing one can also be covered.
- The surveyed period is long enough to include a mix of old projects carried over from the previous CP, that should have produced ample results documented through end-of-project evaluations, and newer projects started during the ongoing CP.

Figure 1: Timing of a CP Evaluation



The main disadvantage of a long evaluated period is that an additional effort needs to be made to locate and interview staff members and government officials who were in charge at the end of the last CP and involved in the design of the current one, but have since changed jobs. But this technique - systematically used in WB and IADB – also helps “getting the context right” and knowing precisely what those people wanted to achieve with the evaluated CP. It is to be expected that new teams and managers would criticize or take some distance with the work of their predecessors, so it is only fair to ask them their side of the story. In other words, looking deeper in the past has a cost but it may build up institutional memory and bridge some of the information gaps created by high staff revolving rates.

The timing of the evaluation itself should be shortly after mid-term. CPEs take time, at least six months (see 5.8 below), so it is probably better to start early.

⁸ OECD/DAC, Op. Cit., p. 23

5.2. Finding a Focus

CPEs are expensive evaluation exercises. As currently practiced in UN agencies and IFIs, and depending on its scope, a country review or evaluation can cost between \$100,000 and \$500,000, and requires from 3 to 15 man.month. However, even the World Bank, whose CAR are quite resource-intensive, call for caution when defining the scope of the exercise. CPEs lend themselves to an all-embracing *scope*, often at the expense of *depth*. An average CP is made of two to three dozens projects and a number of non-project activities (advisory/advocacy role, coordination, etc.). Evaluating them *all* at some reasonable depth would in most case not be advisable.

And yet evaluators have to study at least part of the CP in some depth. If because of time constraints they cannot go beyond donors and government offices in the capital city and collect new data from the field, then their evaluation would be of little value-added to the government and the country office. It would tell them things they already know. Narrowing a CPE focus is therefore almost always a necessity. Some agencies discourage the analysis of project-level data, and keep the focus strictly on CP-level issues (how was the CP designed? is it relevant? etc.) rather than looking down at the nitty-gritty of implementation. Others invest whatever resources it takes to be as comprehensive as possible in a few carefully selected countries where they perceive a need for a full-blown review (WB, and probably UNDP in the near future), and perform lighter reviews in other countries.

A common and workable solution to this problem is to draw a sample of projects and non-project activities that will be evaluated. Such a sampling is more often done implicitly, based on the perceived importance of projects from a financial or strategic point of view, than explicitly, i.e. with a list of the entire portfolio to which one applies documented selection criteria.

Another way to improve CPE focus is at the level of the TORs, by selecting a concise evaluation framework or problematique. That is to say a few inter-connected themes and issues that are believed to be relevant and pertinent, and should be analyzed consistently throughout the CP's most important projects. In most cases the evaluation framework can be constructed around the core assumptions or strategies of the CP. For instance, if a given CP goal were to contribute to poverty alleviation via a number of strategies, then it would make sense to test the effectiveness of each strategy as implemented in a sample of projects or non-projects activities (particularly if the coming CP is expected to pursue similar goals, see below).

In short, a CPE has to be selective in the range of issues and projects it can cover.

5.3. How to maximize CPE usefulness for the next CP?

One commonly encountered challenge is to make the study of past aid projects useful for the planning and implementation of future aid projects. The issue is linked to the operational objective of CPEs (help plan and implement the next CP). Among central evaluation offices, the accountability objective (report to the Executive Board on CP implementation) has sometimes been seen as more important, but even among Board Members' interest would probably run higher for a forward-looking evaluation than for a backward-looking one. The submission of a given CPE to the Board is also a way to help the Board make an informed decision on the next CP.

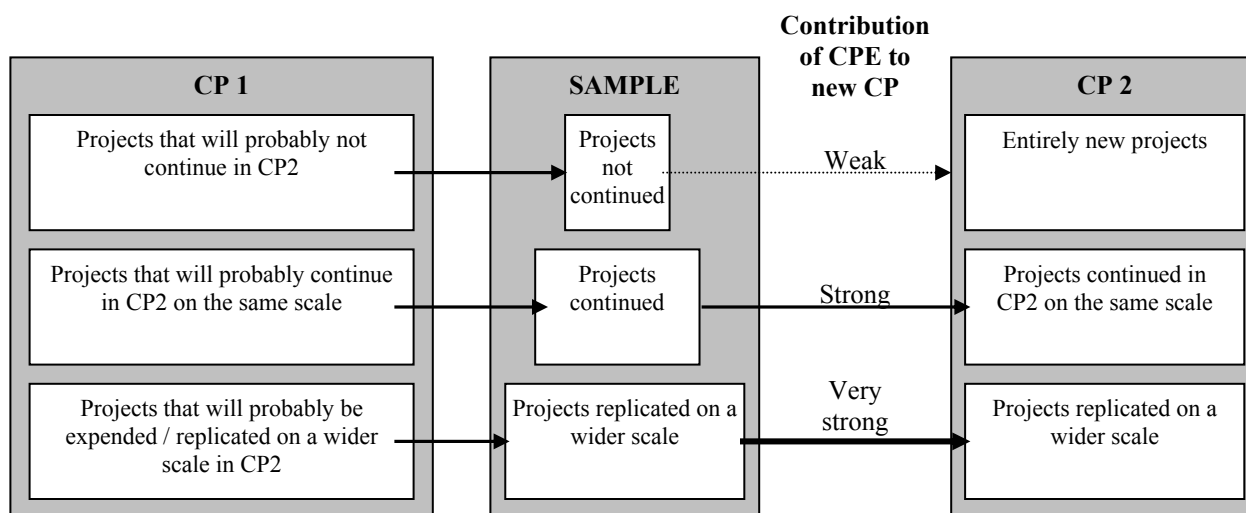
There are a number of reasons why lessons from the past might not apply to the future, the main one being that development situations, challenges and opportunities change over time, whether as a result of programme activities or not. Similarly, the mandate and corporate objectives of UN organizations change overtime. Subjected to the combined pressures of a changing mandate and changing local conditions, new CPs tend to be quite different from past ones in terms of the types of projects being funded, the mix of strategies being used, etc. Whether they are successful or not, projects and strategies do not last forever. Evaluators looking at the performance and results of past projects are often in the position of making the "autopsy of a corpse" rather than diagnosing a condition and prescribing a cure for a "live subject" (i.e. a project or strategy being continued during the next CP).

In practice, the problem is not as serious as it appears. There is often more continuity than meet the eye between two subsequent CPs. Typically, change is more pronounced in the CP as written than as implemented, partly because of change rhetoric, and partly because change is often met with resistance and avoidance. In spite of new terminology, changing mandates and the desire of fast-revolving staff and ministers to impress their mark on a CP, habits and cultures die hard. Cooperation with key local partners is likely to change at a slow pace, even though they receive assistance through new projects designed with new concerns in mind.

When a CPE is conducted – i.e. typically shortly before the new CP is designed – the field office management would usually have an idea of which projects are going to be terminated, replaced by other projects/approaches, and/or extended during the next CP. It may also have some notions of entirely new projects (e.g. as a way to implement a new theme or approach in the organization's mandate). One way to maximize the evaluation contribution to planning is to agree before hand with the field office on the evaluation focus. As explained in previous sections, a CPE usually cannot cover the entire country programme in any great depth. It is therefore necessary to select a sample of projects and/or approaches to study. The sample is best selected in consultation with the field office. It should first and foremost focus on projects and approaches that are going to stay or will be replicated ("live projects"). At the same time, it is important that the sample remain roughly representative of the entire CP. It is advisable to include a small number of abandoned projects in the sample, particularly when:

- there might be a large number of abandoned projects;
- some abandoned projects absorbed significant resources (accountability);
- some abandoned project(s) performed particularly well (success stories);
- some abandoned project(s) are expected to produce lessons of interest for the pipeline (e.g. unsuccessful approaches to avoid in the future).

Similarly, a concise evaluation framework (TOR issues) negotiated with the CO and centering on key strategies pursued during the ongoing CP and likely to be continued would help focus the CPE on questions that are really relevant both for the evaluation of the past CP and the planning of the next one. In short, **the potential orientations of the next CP should be taken into account when drawing the CPE sample and selecting TORs issues.**

Figure 2: Sampling as a Way to Maximize CPE Usefulness

5.4. External vs. Internal Evaluators

Most agencies rely on external evaluator to conduct CPEs, usually for reasons of greater objectivity and credibility vis-à-vis the Board. However, it appears that the level of ownership or “buy-in” from country management teams has been a recurrent issue in many agencies. The consultant is aware of quite a number of cases where operational units and COs rejected CPE findings and recommendations altogether. CPEs are certainly viewed as more important and perhaps more threatening to Country Offices than classic project evaluations, because they look at how the CO is managing the overall Country Programme. In contrast, project evaluations are confined to small segments of a CP and criticisms made in this context are likely to be perceived as less threatening to the Country Office.

The best practice to ensure local buy-in seem to include the following:

- Hire external evaluators to constitute the core of the evaluation team, and enroll CO staff members and Government representatives in the team (preferably staff with an M&E role) as real participants, with the responsibility to write parts of the report;
- The process should be viewed as fair. A systematic, standardized process applied across countries is more likely to be perceived as fair than a more loose and intuitive process;
- Involve the CO in every step of the evaluation (desk review, TORs, project sampling, feedback on results, etc.) for example through a small evaluation management team composed of key CO staff;
- Obviously, the evaluation team should present its preliminary findings and recommendations to a debriefing meeting in country, with the CO, governments and other stakeholders participation as appropriate. Feedback obtained in the debriefing meeting should be reflected in the final report.

5.5. Interviewing Beneficiaries in CPEs

As explained above, it is important, in order to reach independent, credible conclusions, that the CP evaluators collect data from a broad range of stakeholder, from programme beneficiaries (communities, SCOs, government officials, etc.) to project implementing partners and associated donors in country. Experience shows that evaluations which rely on solely information already documented or collected from the CO, the Government and programme staff tend to lack objectivity. Independent, external evaluators

also have a comparative advantage over programme staff when it comes to eliciting feedback from other donors, beneficiaries and implementing partners.

The problem is that 360° stakeholders interviews are time-consuming, so it is important to remain focused. This approach should probably be limited to a handful of important projects that have already produced significant results.

Beneficiaries are probably the most important stakeholders group as far as results (appropriateness, quality, quantity, equitable access to services, etc.) and impact assessment are concerned. The use of formal, already existing groups (community organizations, cooperatives...), rather than ad-hoc ones, tends to increase the chances for panel legitimacy and capacity to articulate criticism.

While some agencies require their evaluators to undertake some fieldwork themselves, some other agencies have experimented with formal surveys and focus-group interviews performed by local firms / NGOs either before or during the CPE process. Whatever the technique used, the aims of focus group and beneficiaries interviews in CPEs should be to:

- cross-check information collected from programme staff;
- collect interviewees' views on programme/project results, quality and usefulness;
- collect interviewees' views on programme/project problems, and suggestions to overcome them;
- collect suggestions on what type of activities or strategies would be most useful for the next CP.

5.6. Looking for Impact

In terms of CPEs contribution to impact assessment, UN agencies are holding modest expectations. The general view is that it is too difficult in an evaluation of such a broad scope to dig deep into each activity and gather sufficient evidence to make impact statements. Obviously, it would be very difficult in the context of a country review or evaluation to gather primary data of the sort and quality necessary to “prove” an impact statement scientifically. This is research, and it takes time. Whatever contribution to impact assessment a CPE may have, it is likely to be through:

- the aggregation and interpretation of previously collected quantitative and qualitative data
- the collection of opinions and ideas of selected stakeholders of selected projects

The first point above is worth mentioning because CPEs could potentially become a data-gathering cluster or echelon in a system-wide information management system. Collecting already documented impact statements (from evaluation reports, studies, etc.) and discussing their merit based on the evidence provided is a worthy task in itself. Some agencies however mentioned the lack of data as a reason for not assessing impact, which would indicate that many of their project monitoring systems and evaluations do not provide sufficient or credible impact information.

In 2000, the UNDP evaluation office experimented with a methodology for Country Level Impact Assessment (CLIA) that was doing precisely that: collect impact testimonies from all stakeholders of a few sampled projects, compile already documented impact-related information, and try and reach some conclusions in the form of documented impact assumptions. One of the rare agencies looking at impact in their CPEs, the WB is applying more or less the same approach, i.e. review reports present impact data collected during the CAS through regular M&E systems rather than new data collected as part of the review.

CPEs should not attempt to collect primary impact data other than stakeholders' opinion. Regular project-level M&E practices and surveys should be used to gather information on final results of

external assistance at the country level. This data can then be pulled together periodically through CPEs to pass a judgment on the impact of at least the largest projects in the CP.

5.7. Blending Review and Evaluation

CPEs have often suffered from the comparison with annual, mid-term or final reviews. What the term “review” refers to is essentially a mandatory monitoring procedure, an operational process oriented towards implementation rates, monitoring and planning. Of course, the term “review” is vague-enough to include more evaluative exercises looking for relevance, partnerships, long-term results and cost-effectiveness, such as the WB CARs. The idea of staging two distinct, intensive and expensive CP M&E exercises – one called a review and the other an evaluation – appears to be a dead-end. Cost-effectiveness alone would call for a single CP-wide M&E exercise per cycle. Beside, there is a certain conceptual overlap between monitoring and evaluation. It is a well-established principle that monitoring and evaluation can and must interact and benefit from each other in practice.

The alternate path of developing a unique, mandatory M&E exercise blending monitoring and evaluation appears more practical. It has so far been followed by the WB, IADB, UNICEF and UNDP (sometimes after experimented with the previous option). Applying the format in a flexible manner is key to the success of this option. In principle, the monitoring or “review” aspect should be quite comprehensive, documenting portfolio implementation and results in a systematic way. But the evaluation part does not need to be comprehensive. On the contrary, it must remain focused on a handful of evaluation questions and, depending on resources available, cover only a sample of sectors and/or projects.

Hence the idea of various sizes and shapes for CPEs (or CPRs) pursued by the WB and UNDP. The size and depth of the review depends upon the country and chiefly the size of the country programme. Small country programmes are subject to simple reviews of programme implementation, achievements and constraints with no or little evaluative dimension; medium-size CPs are subject to a review with a modest evaluation framework; and large CPs are subject to a “full-blown” review with a more ambitious evaluation agenda framework.

Blending reviews and evaluations boils down to setting an evaluation framework, a set of evaluation questions to complement the monitoring review information, assess the use made of donor resources under the ongoing CP, and help plan for the next one.

In setting an evaluation framework, it is considered good practice for the central evaluation office – or whoever is “driving” the evaluation exercise – to enter in meaningful discussions with operational units and country level partners to ensure relevance and propriety of the evaluation (buy-in). This connection with operational units needs to be maintained all along the exercise, since there might be a risk of dissociated ownership in this combined review-evaluation mode: dissociation between an “evaluation part” driven by the central evaluation unit and not really considered or owned by the operational units on the one hand, and a “review part” driven and owned by operational units alone (country or regional office). The aim is to maintain dual ownership on both parts all the way through to a unique final document presented to the Board and resulting in effective change at the country level.

In practice, the final document is assembled and submitted to the Board by a unique locus in the organization that is mandated to do so, most often the evaluation office. Yet operational units should have a right to review and comment upon the document before its presentation to the Board. Some donors, including the WB, include in the report a “Management Response” to each and every recommendation explaining the manner in which they are going to be implemented (or not) in the next CP. In the case of the WB, the Management Response is presented to the Board with the CAR report, usually a few weeks before the new CAS.

Blending review and evaluation also entails strong cooperation between operational units and the central evaluation unit.

5.8. How to Use Standard TORs?

Some agencies have tried and developed standard TORs for CP evaluation or reviews (UNICEF, UNDP, WFP). In general, issues listed in standard CPE TORs reflect classic, project-level evaluation frameworks: relevance, efficacy, cost-efficiency, results and sustainability are the major headings. Classic evaluation issues such as design quality, coherence of objectives and quality of partnerships can be applied to the CP level *mutatis mutandis*.

There is nothing wrong with standard TORs in principle. A set of standard evaluation issues applied corporate-wide in all CPEs would in theory allow comparisons to be made between CPEs, and possibly some sort of generalization of findings. It also appears more practical to use standard TORs than to tailor-make them for specific CPEs. But in practice, the idea has often resulted in very long, impractical TORs with a proliferation of issues and questions. By definition, a CP evaluation is more complex than a project evaluation because it includes a greater diversity of interventions and sectors. But it is important not to embrace too much. Experience proves that, faced with impractical TORs, consultants make their own choices and focus on what seems feasible to them, resulting in haphazard or at least non-standard evaluations.

One solution may be to adapt the standard TORs to each specific CP by selecting, within the corporate evaluation framework, a sub-set of issues particularly relevant to this particular CP, its particular sectoral and strategy mix, and the proposed sectoral or strategy focus of the evaluation. This could be done by dividing the corporate TORs into sections that can be tested and used independently (e.g. sections on the Rights of the Child, service delivery, CP design and management issues, etc.). Each section would be tested and refined in a small group of countries, allowing for some level of generalization of findings.

To avoid consultants defining their own focus when faced with very broad TORs, standard TORs should be streamlined and considered as a “menu” of questions. The evaluation office and operational units should make the effort of selecting from the “menu” a short list of issues to define an evaluation agenda particularly relevant to the CP.

5.9. Scheduling the Review

The following is an illustration of how a combined review-evaluation exercise could be scheduled in practice in an organization like UNICEF. The sequence may not reflect UNICEF procedures, but it builds upon current review practices in the organization, and notably the idea of a phased review, with a number of sectoral⁹ reviews preparing a final and overall review meeting.

Tasks	Responsible party	Duration
1. Agree on principle and schedule the review	HQ, CO, Govt	1 week
2. Nomination of review management teams in HQ and in country	HQ, CO, Govt	1 day
3. Desk review of CP quality, focus, objectives and evaluability	HQ	1 week
4. Trace and interview agency staff that were involved in CP design, get the historical context right and describe what the agency really intended to do	HQ	1 week
5. Define a concise evaluation agenda (sectoral focus for in depth evaluation, TORs)	HQ, CO, Govt	2 weeks
6. Define review period, including end of last cycle		
7. Select a few evaluation questions to include in standard review reports (based on evaluation agenda, optional)	HQ	2 days
8. Reporting by projects/programmes and country office (on non-	Projects and CO	4 to 6 weeks

⁹ The CP may be divided in other types of clusters than the *sector* (e.g. *themes* such as poverty alleviation). What is important is to divide up the CP in pieces that make sense for local stakeholders, possibly along CP articulation or CO management lines.

project activities) using standard review reports		
9. Summarizing of project reports into sectoral reports by local consultants	Local consultants	2 weeks
10. Hiring, briefing and arrival of evaluators in country, about one per sector under evaluation focus	HQ	1 week
11. Sectoral review meetings, with participation of CO, government, main implementing partners and evaluators, discuss sectoral reports and propose recommendations for implementation of current CP and focus of next CP	CO, projects, consultants Evaluators	Govt, local & 1 or 2 weeks
12. Sampling of a small number of projects and non-project activities for in-depth evaluation (3 or 4 per sector)	Evaluators	1 or 2 days
13. In-depth evaluation of selected projects and non-project activities by evaluators, through stakeholder interviews including beneficiaries	Evaluators	2 weeks
14. Drafting of sectoral evaluation notes by evaluators	Evaluators	1 week
15. Sectoral evaluation notes communicated to, discussed with and approved by country management team and evaluation office in HQ (if necessary, evaluators can fly to HQ for debriefing)	Evaluators, HQ	CO, 1 or 2 weeks
16. Overall review meeting (3 to 5 days). Presentation of sectoral review reports and recommendations, and of evaluation notes & recommendations. Adoption of a consolidated set of recommendations for current CP and focus of next CP. Devote one or two days to cross-sectoral issues.	CO, projects, Evaluators	Govt, 1 week
17. Evaluators debriefed in HQ	HQ, evaluators	1 day
18. Drafting of final report, pulling together the context, CP goals and evaluability, evaluation notes, sectoral review reports as annexes or tables, recommendations made in final review meeting	HQ, evaluators	4 weeks
19. Communicate draft report to operational units and government, collect comments and management response to recommendations	HQ, CO	2 weeks
20. Finalize report and present it to the Board	HQ, evaluators	1 week
Total:		6 months

The whole process above should take about six months and involve four to five man.month of international consultants plus three to four man.month of national consultants. It is significantly heavier a process than the one used in the 1990's by UNICEF in its first CPE trials (1 or 2 man.month), a duration that was clearly not sufficient.

–

ANNEXE:

PERSONS MET AND INTERVIEWED

Mr. Jean Quesnel	Director, UNICEF Evaluation Office– EPP
Mr. Joseph Foubi	Senior Research and Evaluation Adviser, UNICEF Evaluation Office – EPP
Mr. Michael Constable	Policy Adviser, UNDP – OSG
Mr. Nurual Alam	Deputy Director, UNDP – EO
Ms. Elena Marcelino	Evaluation Specialist, UNDP – EO
Mr. Masa Kato	Director, FAO – PBEE
Ms. Susanne Frueh	WFP – EO
Mr Nils Fostveldt	Operations Adviser, World Bank – OED
Ms. Poonam Gupta	World Bank - OED
Mr. Sixto Aquino	Deputy Director, IADB – OVE
Mr. Jean-Michel Houd	Former Deputy Director, IADB – OVE