

FINDINGS FROM THE EVALUATION

3.1 RESULTS ORIENTATION

3.1.1 EVOLUTION OF STRATEGIC GOALS AND EFFECTS ON THE ORGANIZATION

The goals of an organization define the framework within which programmes and organizational structure are developed. They also provide a reference point for stakeholders and other parties with which the organization has a relationship. This has been a particularly important issue for UNDP, whereby statements about the use of goals have been intended to focus the programme.

Table 5 lays out the evolution of the goal statements of UNDP between the first MYFF and the forthcoming Strategic Plan. Aside from goals 6 and 7, which are related to internal management within UNDP and the United Nations in the first MYFF, the substantive development goals have remained almost the same over the last three strategic documents.

Governance (defined as an enabling environment for sustainable human development under the first MYFF), poverty reduction and the environment are constant across all three documents and periods, with only the language changing to reflect the advent of the MDGs. Gender was dropped as an explicit single goal in the second MYFF and instead became one of six 'drivers of development effectiveness'.

The specific goal on HIV/AIDS found in the second MYFF will be dropped in the new Strategic Plan, but as noted under the Strategic Plan's Frequently Asked Questions:

*'As a co-sponsor of UNAIDS, UNDP will continue to be part of the worldwide effort to respond to the HIV/AIDS epidemic. Based on the established division of labour, UNDP is designated as the lead organization on behalf of the UN system for addressing AIDS as it relates to development, governance, mainstreaming, legislation, human rights and gender.' UNDP's work in executing these responsibilities is now reflected within the Democratic Governance and Poverty Reduction and MDG development focus areas.*⁴⁰

With the exception of the statement to achieve the MDGs, all the goals represent processes or areas of activity, without inherent targets or predefined indicators of performance. This makes progress towards them hard to evaluate.

The statements of goals are supported by areas of work at a lower level called strategic areas in MYFF 1, service lines in MYFF 2, and key result areas in the Strategic Plan. Annex 5 Table 3 presents a comparison across these areas of work. This analysis shows that although the wording has changed and some components of MYFF 1, such as area-based and downstream work, have been dropped, there has been little substantive change to the scope of activities. One major innovation associated with the new draft Strategic Plan is the proposal to build corporate outcome indicators into the results framework, but at the time of writing this has yet to be approved by the Executive Board.

Regarding the use of the strategic frameworks as a communication tool, senior managers in UNDP and partner organizations believe that

40 Available online at http://intra.undp.org/results/strategic_plan_faq.shtml.

Table 5. Key strategy documents and evolution of UNDP goals

Strategy Document	Goals Identified
MYFF 2000-2003	<ol style="list-style-type: none"> 1. To create an enabling environment for sustainable human development 2. To eradicate extreme poverty and reduce substantially overall poverty [WSSD Commitment 2] 3. To protect and regenerate the global environment and natural resources asset base for sustainable human development 4. To achieve gender equality and advance the status of women, especially through their own empowerment 5. To prevent or reduce the incidence of complex emergencies and natural, environmental, technological and other human-induced disasters, and to accelerate the process of sustainable recovery 6. To provide effective UNDP support to the United Nations Agenda for Development 7. To achieve excellence in the management of UNDP operations
MYFF 2004-2007	<ol style="list-style-type: none"> 1. Achieving the MDGs and reducing human poverty 2. Fostering democratic governance 3. Energy and environment for sustainable development 4. Crisis prevention and recovery 5. Responding to HIV/AIDS
Strategic Plan 2008-2011	<ol style="list-style-type: none"> 1. Poverty reduction and the MDGs 2. Democratic governance 3. Crisis prevention and recovery 4. Environment and sustainable development

the frameworks have had a positive effect on the presentation of the programme to stakeholders and third parties.

3.1.2 ALIGNMENT WITH STRATEGIC GOALS AT THE COUNTRY PROGRAMME LEVEL

It is clear that the development of strategic frameworks was intended to help focus the programme, but the extent to which focus meant substantive change, or improved specification, is less clear. Increasing focus implies reducing the range of activities being supported. This raises important questions about how UNDP interacts with government (as the owner of the planned results) and development partners. One purpose behind the drive for focus was to enable UNDP to foster areas of competence and avoid overlap with other UN partners. Few changes were

expected from the first MYFF, as that was largely drawn up around current programmes and designed to make the presentation of the portfolio more coherent and logical. There was more scope with MYFF 2.

Analysis of the five case-study countries reveals some signs of changing programmes, but a disparity exists between perceptions of staff and management and the structures of the country portfolios. According to a recent survey of country offices, the large majority of Resident Coordinators and Deputy Resident Representatives believe the practices and service lines introduced in 2002 have benefited their work by providing greater focus, improving knowledge management and focusing on results. This increased focus has facilitated positioning and advocacy at the

national level, especially by identifying areas of legitimate involvement by UNDP.⁴¹ This view was echoed in the countries visited.

The main evidence in support of increased focus is the steady reduction in the number of outcomes, especially since MYFF 2. For example, the Indonesia programme has reduced outcomes from 28 to 5. But much of the change was driven by structural guidance from headquarters limiting country programme outcomes to 10 or fewer under MYFF 2. In three of the five countries, close inspection of the portfolio and discussions with staff revealed that the underlying areas of support have remained largely unchanged. The exceptions are Zambia, where there has been a clear withdrawal from agriculture, rural development and food security, and employment generation and sustainable livelihoods; and Argentina, where the programme saw a dramatic change of orientation following the economic crisis of 2001-2002.

All of the case-study countries characterized this process as a remapping of projects against the evolving strategic frameworks, rather than substantive change. This process has become increasingly elaborate as the language of objectives has evolved and the framework of results has become steadily more complex.⁴² For example in Egypt, the programme in 1993 was configured under two areas of concentration and nine themes. By 2007, the nine UNDP Egypt outcomes were mapped into a framework of five UNDAF outcomes, four MYFF goals, nine service lines and nine core results. It is only with the advent of CPDs and the annual CPAP that a practical tool has emerged that enables UNDP and national partners to agree on the focus of

results in a way that is coherent with the wider span of UN support at the country level. As more countries develop second-generation UNDAFs, the clarity that the UNDAF/CPD/CPAP structure brings has the potential to have a strong positive effect on both UNDP's drive for focus and national ownership of results.

In the five case-study countries, changes to the composition of the portfolio of projects are more likely to arise from other pressures such as: carry over of projects from previous planning periods, the need and opportunity to mobilize resources (examined in the next section), and relationships with other development partners.⁴³ Most importantly, country office staff have observed that outcomes are broad permissive statements that act as umbrellas in the planning process rather than objectives that drive the choice of projects. A 2007 Country Programme Outcome Evaluation for Egypt concluded that greater focus was needed within the outcome areas, a point echoed in the case study for Zambia. There is clear evidence that countries have responded to strategic guidance by removing outlier projects. For example, in Egypt, projects in support of science, space technology and infrastructure have been terminated.

The extent to which corporate strategic plans have focused the portfolio was examined by the Operations Support Group (OSG) in 2006. That analysis concluded that 87.5 percent of a random sample of projects were fully or substantially aligned to the 2004-2007 MYFF. The 12.5 percent unaligned actually accounted for 35.7 percent of expenditure (a finding that suggests most unaligned projects had substantial infrastructure and procurement components).⁴⁴ However, this

41 UNDP, 'UNDP Management Review Phase II – Report of the Management Review Team', Internal UNDP Report, 2007, para 73.

42 Terminology has, at various times, included some or all of: goals, sub-goals, key result areas, outcomes, practice areas, service lines, and core results.

43 Examples were given in Indonesia and Zambia of circumstances where the World Bank assumed the lead from UNDP in post-tsunami reconstruction and decentralization due to a stronger presence on the ground and larger resources. Other instances include rationalization of projects that are within the mandate of other UN agencies, such as support to counter female genital mutilation in Egypt.

44 'Alignment Analysis', OSG/ExO (AB), 30 Oct 2006.

Box 3. Broad, permissive outcome statement

The Indonesia country office has taken the UNDAF sub-outcomes as its overarching framework. The sub-outcomes are broad and cover the full spectrum of UNDP's previous activities with no influence at all on UNDP's portfolio. For example sub-outcome 1.4 reads: "By 2010, increased opportunities for achieving sustainable livelihoods in the poorest provinces of Indonesia through development and implementation of appropriate participatory policies and programmes." The implication is that, under this outcome, anything to do with policy or a programme that increases opportunities for achieving sustainable development can fit. (Indonesia case study)

analysis raises a question about the quality of outcomes. The statements of goals in Table 5 and the areas of work (service lines etc.) in Annex 5 Table 3, reveal that the wording tends to be phrases that describe activities to be carried out, rather than measurable objectives. They are broad, permissive statements within which a wide range of project activities can be clustered (see Boxes 3 and 4.)

Concern over the quality of outcomes has been noted by the Evaluation Office as part of a review

of M&E practices. It found, inter alia, that most results orientation was predominantly at the project output level and that trying to use outcomes as the focus of evaluation raises difficulties of linking the status of outcomes to interventions and the relevance of current approaches to strategic and operational planning in country offices.⁴⁵ The study was able to cite some instances of good practice at the country level and similar examples were also found in the case-study countries.⁴⁶

Box 4. Making outcomes evaluable

An example of training material from the European Commission, which uses a similar definition of outcomes as UNDP, illustrates how careful specification can help planners develop more measurable outcomes.

UNDP definition

Outcomes are actual or intended changes in development conditions that UNDP interventions are seeking to support. They describe a change in development conditions between the completion of outputs and the achievement of impact. For example:

- Improved national capacity to monitor human and income poverty and inequality
- Increased access of the poor to finance (formal, informal, micro)
- Reduction in the level of domestic violence against women

European Commission training guidance

The *outcome* is the effect that project outputs will have on the beneficiary, institution or system. It defines the project's success. Test the outcome by asking is it a realistic statement, specific to project outputs, and measurable at reasonable cost. For example:

- Improved performance in the national statistics office for timely production of human and income poverty statistics of verifiable quality to international standards
- Increased utilization of formal, informal and micro finance by poor households in (designated) area or for (designated) purpose
- Reduction in the (reported or estimated) level of domestic violence against women in designated areas

45 UNDP, 'An Assessment of UNDP Practice at Monitoring and Evaluation at the Country Level', UNDP Evaluation Office, New York, NY, February 2005, para 4.19, 4.21 and Summary.

46 See for example, Table 7 in the Egypt Case Study, which provides an illustration of improving outcome and indicators, UNDP Evaluation Office Internal Report.

The five study countries present a wide spectrum of funding status.⁴⁷ At one extreme, Argentina and Egypt have very low levels of core funding (less than 1 percent and less than 4 percent respectively) and rely on mobilizing funds from donors and the government for virtually all their programmes. The proportion of core funds was higher in Indonesia (13 percent), Moldova (38 percent) and Zambia (65 percent). However, evidence from the five countries was consistent that resource mobilization and delivery were the major internal criteria used to assess overall performance. A key issue therefore is the strategies to use core funds and to mobilize funds in support of outcomes. Table 6 summarizes the explanations given by the five countries for how core funding is managed. None of the countries incorporate results in their decision process.

Mobilization of non-core funds creates a potential conflict for country programme management. Many countries are dependent on fund mobilization to generate a ‘critical mass’ of programme activities and staff their office. As a result, mobilization tends to be a driving force behind the work of programme staff.⁴⁸ Examples exist, such as in the Indonesia office, where there has been success in matching non-core funds to their outcome objectives. But more

often, acquisition of resources is opportunistic and reflects availability of funds from donors and the government rather than being driven by programme outcome objectives.

The 2007 Country Programme Evaluation of Egypt concludes:

“The current partnership strategy is strongly driven by resource mobilization ends and not the other way around. Diversification of UNDP positioning strategy can hurt the total size of the profile in terms of resources, but is likely to deepen impact of UNDP interventions at the policy level towards achieving the MDGs especially forging alliances around difficult but common platforms in governance areas. There is the temptation to invest in areas that carry prospects of non-core resources, resulting in UNDP spreading itself too thin at the expense of focusing on its strategic strengths and expertise. Projects in the pipeline need to be re-evaluated with this caveat in mind.”⁴⁹

This same issue was raised by the Evaluation Office in 2006:

“Conflict of interest and confusion of roles may arise where UNDP seeks to combine the roles of

Table 6. Core fund allocation strategy					
Strategy for Core Funds	Argentina	Egypt	Indonesia	Moldova	Zambia
Use as seed money		✓			
Allocate according to needs				✓	✓
Allocate according to fund-raising potential			✓		
No formal strategy	✓				

47 Funding analysis reflects experience over a variety of years as available at the CO: Argentina 2005-2008; Egypt 2002-2006; Indonesia 2001-2005; Moldova 2002-2006; and Zambia 2004-2006.

48 55 percent of staff agreed with the statement “Because most of our funds are raised through cost sharing or from donors, we have little scope in allocating resources across our programme or within outcome areas according to results;” 62 percent agreed with the statement “In my office, country programme staff are under more pressure to raise resources and ensure timely delivery than on enhancing the contribution by UNDP to achievement of the outcomes.”

49 Afaf Abu-Hasabo et al., ‘Evaluation of Egypt Country Programme (2002 – 2006)’, UNDP Egypt, March 2007.

*policy coordinator, donor, rights advocate, neutral broker and project implementer. In particular, areas where UNDP has a comparative advantage but is unlikely to mobilize external resources—such as donor coordination, neutral broker activities and advocacy for human rights—can be crowded out by activities for which cost-sharing assistance is more readily available and government consent easier to secure.*⁵⁰

Reforms to human resources management over the period were driven by the Administrator's Business Plans and the aim to reorient UNDP upstream to policy engagement rather than processing of projects.

The experiences in the five countries share some common threads, but are largely different. Changes in Argentina were linked to the expansion in the country programme. These changes did enable some realignment of staff to substantive programme areas but were not linked to results-based management as such. The organizational structure of the Egypt Country Office was changed by each of the previous two Resident Coordinators and is currently being reassessed by the new incumbent. Unique among the five study countries, Egypt adopted a flat structure in 2002 with responsibilities for projects allocated to individuals without a structure based around programmes or outputs. Various systems were put in place to manage for outcomes, but these have not been maintained and staff feel this has contributed to a lack of strategic direction during the transition to the present management team.

Three of the five country offices underwent reprofiling exercises with varying degrees of reorganization between operations and programming staff, including the development of new job descriptions and the need for new skills. Indonesia also adopted a flat structure but with units grouped around programme areas and staffed according to programme size. Zambia reported missing an opportunity during reprofiling, with existing staff being fitted into new job

descriptions rather than new skills being found or developed. Moldova underwent reprofiling and has continued with further reform, culminating in a change process and restructuring in 2006. This is the nearest example of a results orientation, with the drivers being:

- Aligning staff around delivery against key UNDAF outcome areas relevant to UNDP
- Addressing the burdens imposed by the rapid increase in programme resourcing
- Strengthening UNDP's resource mobilization capacity⁵¹

The RCA has been adopted in all offices with varying degrees of success. It is generally seen as strengthening the potential for accountability with the link between staff member to supervisor to senior manager. RCAs do not yet include measures of whether staff are managing for outcomes. For this reason, the Zambia office has found the RCA to be a more powerful tool for operations staff, whose procedures foster stronger accountability and feedback, than programme staff where the focus should be on managing for outcomes.

Specific criticisms were raised about three aspects of the RCA, both in the country offices and in headquarters:

- First, although staff performance assessment is rated on a five-point scale, in practice only three points are used, as the lower scores of four or five trigger management action, which all parties seek to avoid. This results in clustering of staff and poor differentiation in assessment.
- Second, there is no difference in the financial reward for good performance between grade three and the highest grade of one, which reduces incentive for higher performance.
- Third, the targets set for performance assessment are self selected and, in an estimated 70 percent of cases, are agreed retrospectively during the year of assessment.

50 'Annual Report of the Administrator on Evaluation in 2005', DP/2006/27, para 67.

51 UNDP, 'Moldova Case Study Report', UNDP Evaluation Office, Internal Report, para. 4.33.

Key points

- UNDP's goals in the strategic frameworks have changed in presentation but the underlying areas of work have remained almost the same.
- The focus areas under the goals have been rationalized and simplified, but it is hard to identify substantive change to the scope of activities at country level.
- Managers and staff in country offices believe the MYFFs have helped to bring focus and improved positioning and advocacy. They have helped remove outliers on country portfolios, but have otherwise had little effect. Projects have just been mapped to the new frameworks.
- Outcome statements for country programmes tend to be broad and permissive rather than providing a test for how well a project contributes.
- Mobilization of resources is, in practice, a more powerful driver of individual performance among programme staff than achievement of results.
- There is some evidence that country offices organize their staff for delivery of results.
- The RCA is welcomed as a tool but in its current modality does not support managing for results.

3.2 MANAGING FOR RESULTS

3.2.1 MONITORING AT THE COUNTRY PROGRAMME LEVEL

Monitoring at the country level is an interesting discussion. To a large extent, the expectations set out in Table 4 of Annex 5 have not been achieved, but the five case-study countries reveal some local initiatives as country managers try to grapple with the challenge.

Table 4 in Annex 5 brings together responses from the case-study countries about progress towards outcome monitoring. The findings show that progress is uneven, but there are consistent signs of greater use of baseline data and setting indicators for outcomes. There is also evidence of progress towards more collaborative work with partners and greater use of national surveys and tools, such as the preparation of the National Human Development Report (NHDR). The weakest areas are the intermediate steps that demonstrate how UNDP-supported projects are delivering contributions towards outcome objectives. This type of analysis will not be derived from national statistics and requires both a clear causal pathway against which progress can be charted and data collection that is specific to

project activities. Similarly, the country visits reveal little progress towards data collection about perceptions of change among stakeholders and systematic qualitative information.

The shift to outcome monitoring and the 2002 monitoring policy has contributed to a growth of local initiatives as managers look for ways to improve their control of the project cycle. These initiatives are particularly strong in those countries that have invested in dedicated M&E staff or M&E units: Argentina, Egypt and Indonesia are examples in the study sample.⁵²

- In Argentina, there has been a focus on developing tools for the whole project cycle, with a strong focus on monitoring and risk assessment. A new ex-ante evaluation procedure has been introduced with procedures such as review by the Local Project Approval Committee.⁵³ Some of these procedures are being taken up by the Regional Bureau for Latin America and the Caribbean (RBLAC) as part of regional guidelines.
- The Egypt country office has retained the long-standing quarterly and annual project reports, joint annual tripartite reviews and

52 Currently, 25 country offices have M&E specialists and 10 have a dedicated M&E Unit.

53 The ex-ante evaluation procedure is referred to elsewhere in this report as 'appraisal' following normal project cycle convention.

project self-evaluations that were made non-obligatory in the 2002 policy. Standard approaches to M&E have been introduced in project documents and the office reports M&E becoming stronger—borne out by a review of how outcomes and indicators for programmes have improved over the period. The office still finds it hard to link project monitoring to programme outcomes for the ROAR.

- Indonesia has developed a number of tools, most prominent of which is Touchstone, a locally developed project and programme management software. A database brings together information to demonstrate the way in which individual projects contribute to programme outcomes. (See Box 5 for a range of instruments in use.)
- Less progress has been made towards monitoring in the Moldova programme. The national setting was formerly not very conducive—national planning only started in 2002 with the uptake of Medium Term

Expenditure Framework and few government systems are in operation. The country office monitors projects through project reports and annual programme reviews, though stakeholders report it is difficult to get government counterparts to devote the time necessary for in-depth discussion of issues.

- Zambia appears to have made most progress towards working through government systems and with the United Nations and other development partners. This is a reflection of the country having a large budget support programme for aid. Harmonized working was fostered by a joint assistance strategy aligned with the Fifth National Development Plan (which has an M&E component). UNDP's programme document is directly aligned to the plan as well. The programme works through sector advisory groups and participates in joint programme annual reviews.

A common complaint is the absence of tools to relate progress with projects to contributions to

Box 5. Monitoring and evaluation system components

The Indonesia Country Office's internal M&E functions are supported by the appointment of M&E focal points in each programme unit who interact with the Planning, Monitoring and Evaluation Unit in bi-weekly coordination meetings.

In addition to the standard corporate tools such as MYFF, the ROAR, and outcome evaluations, a range of other instruments for monitoring are operational. These include:

- Touchstone (project management guide)
- Project Data Base System (has the format of ROAR, tracks results quarterly)
- A five-year M&E Plan
- Unit/programme workplans
- Unit-level M&E Guides (e.g., Crisis Prevention and Recovery Unit and Environment)
- Quarterly review meetings with the government
- Bi-monthly meetings with the State Secretariat, the National Development Planning Agency and the Ministry for Foreign Affairs
- Joint field visits and joint monitoring of selected projects with government counterparts
- Management team meetings
- Unit head team meetings
- Bi-weekly meetings with M&E focal points
- Annual Results Reports
- Reporting on progress towards the achievement of MYFF and balanced scorecard targets

programme outcomes. This disconnect contributes to the finding that more progress has been made with reporting than in the use of reports. One consistent message from the country case studies is that the ROAR is primarily a reporting tool to headquarters, with little operational value at country level. The shift from projects to the ROAR undermined the evidence about results in UNDP projects and failed to provide new tools to track progress from projects towards outcomes. Staff were questioned about the ROAR in the staff survey on results-based management. This produced the counter-intuitive finding that 62 percent of staff considered the ROAR to be an effective outcome monitoring tool, but among managers (Deputy Resident Representatives and above), 59 percent think it is not effective—a finding more in line with observations in the case-study countries.

Another channel of performance reporting from countries is the Resident Coordinator’s annual report.⁵⁴ This is described in the current guidelines as “an essential element of the accountability and results framework for UN operational activities.” The report follows a structured format with some narrative and tabular data. It brings together information about performance during the past year in the following format: an introductory letter highlighting progress in UN reform, major results achieved, good practices and lessons learned; results and use of funds table; UN Country Team (UNCT) workplan matrix; and good and bad coordination practices and procedures. A sample of these reports was examined from each of the case-study countries and an example of reporting against one UNDAF outcome is reproduced in Table 5 of Annex 5. The report covers key issues of performance, but the information is limited to observations about delivery of outputs, with little or no dialogue about contribution to outcomes. The nature of the presentation differs between countries, and reveals the varied quality of outcomes and indicators already noted in this

report. The example in Annex 5 illustrates how difficult it is for a reader to track what progress is being made and to understand how the UNCT will contribute to development outcomes.

3.2.2 ADJUSTMENT OF RESOURCES

Adjusting work in response to results is the cornerstone of an effective results-based management system. This study has failed to find any convincing evidence that suggests that results are influencing management. All country offices recognized that adjustments are made to projects from time to time, but there were no obvious examples of information about results guiding decisions about finance or human resources, especially not at the level of programme outcomes. Some practical examples were noted in Moldova:

- There has been a balance between support at the central and local government levels. The balance between these two streams of work has been a continued focus of the programme, affected by UNDP’s assessment of the prospects of creating sustainable capacity at the central level.
- In its strengthening of the government’s poverty-monitoring capacity, UNDP initially gap-filled this function by directly contracting and managing consultants. UNDP realized that this was unsustainable and used the 2005 outcome evaluation to help convince the government to start creating its own capacity in this area. UNDP supported this process by building the capacity of the new staff.
- The Moldova programme is closing its project supporting strengthening of the Parliament based on evidence that support is unlikely to be effective, since there is insufficient support from the partner.
- In its project supporting re-integration of victims of human trafficking, monitoring highlighted that sustained impact depended upon helping women find jobs. Therefore, the subsequent phase of the project has focused upon this challenge.

54 Copies dating back to 2002 can be found via the UNDP internal website at: <http://www.undg.org/index.cfm?P=490>.

These examples illustrate practical responses to information from a range of sources. The formal results-based management tools may or may not have been prominent, but managers have to make decisions with whatever information is on hand—in this case, a mixture of monitoring and periodic evaluations.

3.2.3 RELATIONS WITH THE REGIONAL BUREAUX AND SUBSTANTIVE CORPORATE OVERSIGHT

Despite the intention in the mid 1990s for regional bureaux to undertake holistic oversight of country office performance, the management relationship between country offices and headquarters has remained ill-defined.⁵⁵

Most significantly, the regional bureaux have not been considered accountable for the development effectiveness of country operations in their region. One of the reasons for this is the intention to decentralize management and accountability to the country level and the limited number of entry points at which regional bureaux can intervene in country programmes. Projects are planned and approved within country offices, so one of the main entry points for the regional bureaux is the CPD approval process, which occurs every four years. Even this is of limited scope as the CPD is, on paper at least, a government programme presented to the Executive Board by UNDP.

In practice, the regions have approached their role of substantive oversight in a variety of ways, some more imaginative than others.⁵⁶ Interaction was more likely to occur around operational issues, and good performance was traditionally seen as resource mobilization and delivery. Indonesia found the interaction with Regional Bureau for Asia and the Pacific (RBAP) over the development of the 2006 CPD to be ‘timely and conscientious’, although most interaction was about process and procedural compliance rather than substantive content.

All the regions have had to adjust their relationships to accommodate the new monitoring tools: ATLAS, the balanced scorecard, Global Staff Survey, RCA and the Dashboard have changed the information environment for regional directors. These tools predominantly deal with financial and process information, and that predominantly at the level of projects.

The Moldova and Regional Bureau for Europe and the CIS (RBEC) experience is interesting. In response to ambiguity over the precise meaning of oversight, RBEC has developed its own tool to maintain oversight of country-level operations and effectiveness. This is the Strategic Note, which was introduced in 2002. The Note is a six-page document, drafted annually, that briefly sets out progress on resource and programme issues in the past year and major actions in the coming year. More importantly, the Strategic Note includes mutually agreed targets on what the country office will deliver in the coming year and what has been delivered against the targets agreed for the past year. Targets set for Moldova between 2004 and 2007 are in Table 7 and show that the dominant focus of oversight is on resource mobilization and delivery.

The Moldova Resident Representative reports that there has been no substantive discussion between the Moldova country office and the RBEC over development results. Instead, discussion focuses on resource mobilization and delivery. In common with annual targets set in the balanced scorecard and ROAR, targets need only be agreed upon in April-June, four to six months into the planning year, which considerably weakens their value as tools that drive intended behaviour.

Regional Bureau for Latin America and the Caribbean (RBLAC) and Regional Bureau for Africa (RBA) have developed a Composite

55 UNDP, ‘Annual Report of the Administrator and Related Matters’, 1997, DP/1997/16/Add.7. Addendum: Change Management: UNDP 2001.

56 Definition of oversight has now been given in the 2006 Results Management Guide.

Table 7. Targets in the Moldova Strategic Notes (2004-2007)

Year	Targets
2004	<ul style="list-style-type: none"> ■ Delivery: 100% of USD 2.66 million (of which TRAC is USD 1.48 million) ■ Resources mobilized: USD 1.48 million
2005	<ul style="list-style-type: none"> ■ Core delivery: 100% of USD 1.77 million ■ Non-core delivery: USD 4.2 million ■ CCA and UNDAF finalized ■ A national monitoring system in place, with DevInfo at its heart, helping to track progress towards Poverty Reduction Strategy Papers and MDG targets ■ One outcome evaluation/best practices 'knowledge management product' on local development
2006	<ul style="list-style-type: none"> ■ Non-core delivery: USD 3.75 million ■ Resources mobilized: USD 6.0 million ■ One outcome evaluation (Human Rights & Access to Justice) ■ CPD approved by the Executive Board ■ NHDR on quality of economic growth for human development
2007	<ul style="list-style-type: none"> ■ Non-core delivery: USD 5.0 million ■ Resources mobilized: USD 8.0 million ■ 1 programme evaluation conducted (Joint Programme Support to Strategic Policies) ■ 1 study on non-governmental organizations published ■ 1 joint programme launched (Disaggregated Statistical Data)

Performance Ranking index to track country performance. RBA argues that this has to some extent informed allocation of TRAC 2 funds. However, there is a sense of information overload, as much of the data in ATLAS and other tools is project-based and hard to interpret without supporting qualitative explanation.

The tools provide a strong quantitative asset to oversight. Less strong is the management interface to understand qualitative issues. Regional bureaux typically have a desk officer or programme officer responsible for the day-to-day relationships with the country office. Sometimes these relationships work well, but there are a number of weaknesses:

- Manpower resources mean that, in some regions, each desk officer has to deal with as

many as eight countries.

- Travel budgets restrict country visits to, at most, two a year
- Desk officers are typically employed at grades P3 or P4. Resident Representatives will usually be D1, and in the strongly hierarchical culture of UNDP, that grade disparity makes it difficult for the desk officer to raise critical questions. This results in issues being moved up to the level of regional director or deputy director.⁵⁷
- Desk officers do not always have the technical skills to intervene in substantive programming issues.

Tools such as ATLAS are changing the relationship between regional bureaux and country

⁵⁷ The absence of a deputy regional director in RBA is believed to have undermined the oversight role in the recent past and led directly to RCA target setting for RR and DRRs not being completed for 2007.

offices. As project management information expands in content and coverage, the imbalance of greater attention to finance and process rather than results may be reduced. However, much of the progress reporting in ATLAS and the ROAR is still self assessment and focussed on project performance. Validity only gets checked at the

country level once every four years when an expanded audit is carried out.⁵⁸ The management response to the Office of Audit and Performance Review for 2004 for UNDP Zambia shows that the expanded audit by the Office is a useful process for course correction, but some of the recommendations may have been

Box 6. Office of Audit and Performance Review

The Office of Audit and Performance Review works on a decentralized basis through about 45 staff in total. Some 35 country office audits were completed in 2005-2006. Their extended or full-scope audits deal with good management, including results-based management. Staff have not received any specific training in results-based management per se. The results-based management guide is used as their audit criteria.

The approach adopted is to look at outcome coherence with UNDAF, outcome specification and relation to outputs, targets, and measurement approaches. They make a compliance check to see if the annual review and update of projects has been done and followed up on.

MYFF reports (the ROAR) are examined and their quality, consistency with previous statements and expectations, and the evidence behind statements made are assessed.

In Office of Audit and Performance Review's judgement, results planning and measurement are often of poor quality: output/outcome distinctions are not clear, results are not measurable, and indicators are poor. They believe there is a need to have a procedure for outcome verification.

Key points

- Some progress has been made towards outcome monitoring, and country offices have shown interesting initiatives, but there is little explanation of how projects are contributing to programme outcomes.
- The absence of firm procedures has led to creative diversity in M&E, especially where the country office has a dedicated staff member.
- The ROAR is primarily an upward reporting tool with little utility in the country office or for regional management.
- There is no clear and convincing evidence that results are being used systematically to inform adjustments to the country portfolios.
- Strong and effective decentralization has been accompanied by an ill-defined role of oversight for the regional bureaux.
- Diverse approaches at the regional level, often drawing on the new information tools, have been implemented for interaction with country programmes. The quantitative information in ATLAS and other tools is predominantly project-focused and not matched by qualitative information that explains performance.
- Desk officers in the regional bureaux are an important link in understanding country performance but are overworked, under-resourced, junior in status and often lack technical skills.
- The dominance of self-reporting is not adequately balanced by the current system of periodic expanded audits.

⁵⁸ Although the Office of Audit and Performance Review carries out full-scope audits every year, each year, a different sample of countries is selected. This works out to a frequency of one full-scope audit per country in every four years, on average.

overtaken by events. Hence the frequency of such ex-post audits, though appropriate from a cost perspective, may not be appropriate to foster strategic managing for results (see Box 6).

3.3 EVALUATION, LEARNING AND ACCOUNTABILITY

3.3.1 ROLE AND USE OF EVALUATION

In 2002, UNDP's formal project-level evaluation requirements were dropped and outcome-level evaluations were introduced. New rules, mandated by UNDP's new Evaluation Policy of 2006, now make it mandatory that country offices fund and carry out outcome evaluations and that there be a follow-up management response for all evaluations.

Among the five case-study countries, outcome evaluations have been carried out in all except Argentina, where thematic 'cluster' evaluations have been conducted. Zambia has retained aspects of the old evaluation practice by conducting mid-term and terminal evaluations of projects, but in the other countries, these only take place at the instigation of a development partner. Egypt was the only country to have undertaken a country programme evaluation.

Resources for evaluations is one of the issues that have arisen. Under funding has led to late implementation. In the case of Egypt, findings were delivered too late to influence the CPD process. In Moldova, no allowance was made for the lack of project-level evaluation as a data resource. As a result, the duration of consultant time was too short to gather evidence. Poor quality of consultants was highlighted in several countries and related to budget constraints.

None of the countries had clear evidence that evaluation findings had influenced programmes.

In Egypt, concerns were raised over unclear accountability for the management response and follow up to evaluations. The closest example of a positive management response was found in Zambia for an audit. The Office of Audit and Performance Review Full Scope Audit of 2004 (report released in 2006), contains a matrix stating management's response.⁵⁹ The response shows concrete suggestions or progress already made by the country office towards adjusting operations on the basis of the recommendations in the audit.⁶⁰ The matrix also indicates the responsible manager, the implementation date, level of priority and the cause.

The findings at the country level show that country offices are adapting to the new evaluation policy. In some respects, it is too early to pass judgement on issues of accountability and management response to evaluation findings. But viewed within the wider context of the move away from project evaluation and the overall focus of the Evaluation Office at Headquarters, the evidence suggests a decline in the quality and depth of the evaluation base to support managing for results.

The UNDP Evaluation Office was peer reviewed under the auspices of the OECD-DAC Network on Development Evaluation in 2005.⁶¹ Their findings were generally positive and noted the difficulties of evaluation as UNDP moves into softer areas of policy advice. However the review highlighted some critical issues:

- Evaluability of programmes has been severely and consistently constrained by the performance within UNDP's results-based management systems (para 92 et seq)
- Follow-up systems to evaluations were found to be weak (para 144)
- The relative emphasis has been greater on learning as opposed to accountability (page 32)

59 Office of Audit and Performance Review, 'Full Scope Audit Report: UNDP Country Office in Zambia', 2006, p 30.

60 The response section also allows management to explain/defend progress made vis-à-vis challenges encountered.

61 OECD-DAC, 'Peer Review: UNDP Evaluation Office', December 2005. Conducted as part of Peer Assessment of Evaluation in Multilateral Organizations by OECD-DAC.

Recommendations made to strengthen the role of evaluation as a 'reality check' for the programme and to strengthen evaluation against intended results have been reflected in the Evaluation Policy endorsed by the Executive Board in June 2006. Specific actions include the commitment to issue management responses for all evaluations, the Evaluation Office's maintenance of a system to track management responses and report cases where there are concerns over implementation of the management commitments to senior management, and adoption of a policy of disclosure for all evaluations.

3.3.2 LEARNING AND TECHNICAL SUPPORT

Learning and technical support are supposed to be provided through interaction between country office staff and the practice forums and networks. The practice forums are not only for communication but also have a wider range of organizational functions including knowledge management, advocacy, partnership building, learning and professional development, and providing programme countries access to substantive resources (such as rosters of experts and co-financing possibilities).

Together with the knowledge networks, the SURFs and RSCs were planned to help transform UNDP into a knowledge-based organization. They provide technical and policy advice, referrals, comparative experiences and issues-based applied research from specialists with in-depth, multi-disciplinary knowledge of the regions in which they operate.

Findings from the country offices highlight difficulties the countries have had in putting this support into practice. The value of SURFs' and RSCs' cyclical support in the design of new projects is recognized, but learning from results at the country level is less clearly articulated. The two strongest examples come from Argentina and Indonesia.

The Argentina programme used a cluster evaluation of four projects linked to the Medical Inputs Procurement System to reassess their approach.⁶² One project was subsequently closed, the others were reoriented to new objectives, and the lessons learned were applied to project planning in other provinces. In Indonesia, the 2005 Annual Report describes a process of reducing geographical coverage to focus on provinces with a low human development index and susceptibility to conflict or disasters. The office also cites lesson learning as the force behind a decision to collaborate more with the private sector.

No evidence was found of interaction with the practice areas for lesson learning. Interviews at headquarters with the capacity development, democratic governance and poverty practice areas in BDP reveals a variety of approaches being followed. Provision of advice appears to be demand driven and staff who were interviewed acknowledged that horizontal communication across the practice areas and with other headquarter units, such as the Evaluation Office and the regional bureaux, could be stronger. Too much interaction hinges on personal relationships. A clear tension exists between the view that technical support needs to be responsive to demands and able to cope with a variety of requests from the country offices, and the view that BDP should be more prescriptive about intervention models that demonstrate how to design UNDP-supported actions that lead to programme outcomes.⁶³

The Management Review Team looked at the practice areas in 2006. While the basic approach is in place, implementation of the concept still remains a challenge, as shown in the quote below:

“Having successfully introduced the practice architecture, the next challenge will be to strengthen and refine the practice approach and

62 These procurement projects comprised a significant part of country office resource mobilization at the time.

63 One point of view is that staff in country offices lack the tools to manage for outcomes and, in particular, lack support to plan and implement sound interventions. Some staff argue that only three areas of work have well-developed intervention models with supporting technical advice: capacity development, HIV/AIDS, and support to elections.

its implementation. This is critical, given that country offices, despite their satisfaction with the overall practice architecture, have not received consistent and high quality support across all service lines, and that support has generally not included advice to UNCTs. Concerns have been voiced repeatedly over:

- *The lack of systematic and coherent delivery of policy advisory services;*
- *Inadequate definition of roles and responsibilities, and unclear expectations on part of clients;*
- *A high cost and lack of flexibility in the financial model;*
- *Too broad a scope in terms of themes addressed and products and services offered;*
- *A disconnect between knowledge products and services and business processes; and*
- *Products and services not always adapted to national development context.*⁶⁴

3.3.3 ACCOUNTABILITY

A sound accountability framework would be expected to consist of at least four basic features:⁶⁵

1. Definition of clear roles and responsibilities (accountability relationship)
2. Clear performance expectations and reward systems (transparent incentive mechanism)
3. Credible and timely measurement and reporting of the results achieved (giving account)
4. Mechanism to hold to account (fair review of results, 360-degree feedback, reward achievement or appropriate consequences for under achievement, resolve disputes, apply incentive system, or adjust if necessary)

3.3.4 DEFINITION OF CLEAR ROLES AND RESPONSIBILITIES (ACCOUNTABILITY RELATIONSHIP)

UNDP undertook a global exercise to reprofile country offices to ensure organizational structures and human resources matched the new work flows and processes introduced by results-based management. This was a good attempt to define more clearly roles and responsibilities of country office staff. Three of the five case-study countries participated in this as described earlier. Job descriptions were modified and staff went through a transparent re-recruitment exercise. However, since then, job descriptions of old staff have remained static while new results-based management tools and guidelines have been released, which have added new roles and responsibilities for staff. The main changes in tools include: phased rolling out of ATLAS starting in 2004, together with gradual development of additional features; introduction of PRINCE 2 in 2006; and release of the Results Management User Guide in 2006. These have changed workflows and processes, often with significant impact on staff workload, but with no incentives for use of the tools. The nature of work and type of skills needed has also changed as UNDP refocuses itself on upstream policy and legislative reforms and institutional strengthening. Yet the results-based management procedures and reporting tools take up a large percentage of staff time.⁶⁶

A critical issue in defining responsibilities concerns the focus of development results on project outputs or programme outcomes. Results-based management was accompanied by a shift of attention to outcomes, but this has never been reflected in definition of responsibilities. Staff questioned during the country case studies indicated that there is a high degree of consensus

64 UNDP, 'UNDP Management Review Phase II—Report of the Management Review Team', UNDP Evaluation Office, New York, NY, 2007, para 73. Internal UNDP Report.

65 Early drafts of new proposals for an accountability framework prepared by Bureau of Management were in circulation during this evaluation.

66 The view was expressed that senior staff spend a disproportionate amount of time approving transactions in ATLAS. The main point being that results-based management increased, rather than reduced, the work to be done at country offices, yet many of these country offices have a shrinking allocation of core funds and therefore have to cut down on staff funded by the core budget.

on individual responsibilities and accountability for project outputs, but no consensus and varying views about accountability for outcomes. The staff survey found that 61 percent of staff agreed that roles and responsibilities at all levels are clearly set out and known to staff; 79 percent agreed that the Resident Representative/Country Director is accountable for achievement of country programme outcomes (see Annex 8).

3.3.5 CLEAR PERFORMANCE EXPECTATIONS AND REWARD SYSTEMS (TRANSPARENT INCENTIVE MECHANISM)

At the level of the Resident Representative/Country Director, the balanced scorecard is considered the predominant oversight instrument for setting targets and assessing country office performance. Senior country office management participates in the setting of some of the balanced scorecard targets while others are set by headquarters. The RCA is also at the core of performance evaluation of country office top management. The Global Staff Survey results are consciously monitored by senior management and internal adjustments are made to improve the relationship between staff and management so that the indicators improve. In addition, the partnership survey, together with many other criteria, contributes to performance evaluation of the Resident Representative.

The Resident Representative or Country Director set targets annually, both for their individual RCAs and for the country office balanced scorecard, a process noted above. The balanced scorecard, Global Staff Survey and Global Partnership Survey are increasingly being used for benchmarking country office performance against others in the same region or beyond. Hence there is always a desire by country offices to measure up to, or outdo others in their group.

Below the level of top management, the balanced scorecard is not cascaded down to units within the country office nor is it used for headquarters units such as OSG or Bureau of Management. The main instrument for cascading corporate targets to lower level staff is the RCA, described earlier in this report.

However, the link between corporate targets and individual staff RCA targets, especially for development results, is weakened by the absence of unit-level workplans. The link between corporate and individual staff targets is clearer and stronger for quantitative targets of resource mobilization, delivery and financial accountability. The quality of targets in RCAs gives the impression that UNDP is preoccupied with mobilization and delivery of non-core funds, whilst development results are secondary.

3.3.6 CREDIBLE AND TIMELY MEASUREMENT AND REPORTING OF THE RESULTS ACHIEVED (GIVING ACCOUNT)

In terms of results related to financial and process targets, ATLAS and the Dashboard have enhanced timeliness of measurement and reporting. The inter-linkage of ATLAS and the RCA with the balanced scorecard provides for some transparency in results reporting within the balanced scorecard. The main gap, however, relates to timely and credible measurement and reporting of development results, currently not handled satisfactorily either through the ROAR or through outcome and country-programme evaluation. Without a corresponding pressure from top management for results, the systems tend to exacerbate both a project focus and a reporting culture.

3.3.7 MECHANISM TO HOLD TO ACCOUNT

The end-of-year review of RCA results is participatory, and corporate RCA Guidelines give staff an opportunity to challenge their performance ratings through the Career Review Group. UNDP has three principal mechanisms of holding staff to account: promotion, annual salary adjustment, and contract renewal. Employees funded by core resources have two-year (renewable) contracts, while those paid from extra-budgetary resources usually have one-year contracts.

Top management is aware of the consequences of poor performance through promotion prospects. Employees strive to achieve so that their contracts may be renewed. In theory, these

processes should function adequately. In reality, this accountability framework is weak for a number of reasons.

First, staff contracts and the tenure of office of the Resident Representative/Country Director are both shorter than the period of the CPD and not necessarily in phase with the CPD. Hence accountability for results implicitly concerns only short-term targets—such as resource mobilization, delivery and project outputs—rather than longer-term development outcomes. Since programme staff in many offices are funded from extra-budgetary resources, their accountability for results is aligned more to outputs of projects as opposed to delivering at outcome level.

Second, regional bureaux oversight of the process of measuring and reporting results is weak and hence scores in RCAs, the Global Staff Survey, the balanced scorecard and the ROAR are not immune to strategic manipulation.⁶⁷ The quality and reliability of the ROAR is only independently evaluated at the time of the Full Scope Audit.

3.3.8 ACCOUNTABILITY FOR OUTCOMES

There are several challenges in considering accountability for development outcomes. First, there are influences other than the programme itself, such as other programmes and social and economic factors. Second, many outcomes of interest take a number of years to bring about. There is a need to reconcile the planning and

reporting period with the longer time frames often required for outcomes to occur. Third, UNDP is, in most instances, only a small contributor of resources in comparison to donors and government. So the attribution of outcomes to UNDP's efforts needs to be carefully constructed.

This issue has been examined in other settings and both New Zealand and Canada have concluded that there is a need for a revised concept of accountability to take into account the fact that outcomes are not controlled by managers and that managing for outputs alone is incompatible with results-based management. In both countries, the case has been made that managers need to be accountable for influencing outcomes rather than achieving outcomes, and for adjusting activities and outputs as a result of tracking performance to date.⁶⁸

There is some evidence that similar ideas are being put into practice within UNFPA. UNFPA describes an 'accountability for outcomes' in their 2007 Strategic Plan as—in relation to the above elements—being accountable for:

- Ensuring financial controls
- Achieving and monitoring outputs
- Monitoring outcomes (global trends and outcome indicators)
- Ensuring outputs contribute to outcomes⁶⁹

67 It is possible, for instance, for senior management to persuade staff to rate them favourably in the Global Staff Survey in exchange for better RCA scores, which would also boost his/her RCA/balanced scorecard scores.

68 See Annex 6 references for Baehler (2003) and Auditor General of Canada (2002).

69 UNFPA, 'UNFPA Draft Strategic Plan 2008-2011: Accelerating Progress and National Ownership of the ICPD Programme of Action', Executive Board Informal Meeting, 16 May 2007.

Key points

- The 2002 evaluation guidance has led to outcome-level evaluations and country-programme evaluations. In the countries visited, these were not adequately budgeted and were poorly timed to influence new CPD planning.
- Country evaluations conducted by the Evaluation Office have emphasized learning over accountability and not measured performance against stated intentions.
- Results are rarely used for learning at the country level. The practice architecture in BDP is appreciated for technical support, but has poor linkages with functions such as evaluation and has produced few products clearly tailored to business processes.
- The accountability framework linked to results-based management is weak. Roles and responsibilities are generally clear, but targets are self-selected and poorly linked to incentives.
- Tools such as ATLAS and the balanced scorecard have greatly improved timeliness and access to information but are geared towards resources and process. The ROAR lacks substance on results and is rarely assessed for quality.
- Individuals are tied to a project orientation and accountability for outputs. All accountability for outcomes is vested in the Resident Representative/Country Director.
- Results-based management is about managing for outcomes that are not within management control. Accountability can be set for individual staff to manage for outcomes and this approach has been taken up by UNFPA.
- In the 2006 Evaluation Policy endorsed by the Executive Board, commitments were made to strengthen follow-up systems for evaluations.