



UNEG Study on the Evaluability of the UN Development Assistance Framework

This report was prepared by Paul Balogun, consultant, and was published in December 2006.

This report assesses the evaluability of recent UNDAFs in terms of: a) clarity of intent of the subject to be evaluated (relevance and design of the expected outcome statements and results matrices); b) the existence of sufficient measurable indicators (collection of reliable data for analysis); c) quality of joint monitoring systems; and d) external factors (positive or negative) that have influenced the process and the realization of expected outcomes.

Abbreviations and Acronyms	
AWP	Annual Work Plans
CCA	Common Country Assessment
CP	Country Programme
MDGs	UN Millennium Development Goals
M&E	Monitoring and Evaluation
PRSp	Poverty Reduction Strategy paper
RM	Results Matrix
TCPR	UNGA's Triennial Comprehensive Policy Review
UNCT	UN Country Team
UNDAF	UN Development Assistance Framework
UNDG	UN Development Group
UNDGO	UN Development Group Office
UNEG	United Nations Evaluation Group

Summary

1. The UN Development Assistance Framework (UNDAF), introduced in 1997, aims to support identification and implementation of a collective, coherent and integrated UN system response to national priorities and needs. In 2004, the Guidelines issued by the UN Development Group on the development and use of UNDAFs were revised. As part of this revision, guidance on monitoring and evaluation was included for the first time.

2. This report assesses the evaluability of recent UNDAFs in terms of: a) clarity of intent of the subject to be evaluated (relevance and design of the expected outcome statements and results matrices); b) the existence of sufficient measurable indicators (collection of reliable data for analysis); c) quality of joint monitoring systems; and d) external factors (positive or negative) that have influenced the process and the realization of expected outcomes.

3. The conclusions and recommendations outlined are expected to provide inputs into the update of the UNDAF M&E guidance currently being prepared by UNDG led Working Group on Programming Policy. They are also meant to yield benefits for the on-going process of conducting joint country-level evaluations and to provide inputs to refine the scope of the next round of UNDAF evaluations that will be commissioned by the relevant UNCTs.

4. The assessment is based upon a textual analysis of the 35 UNDAF documents completed in 2004 and 2005, where it can be expected that the 2004 Guidelines would have affected the approach adopted both to programme design and monitoring and evaluation and hence their evaluability. This analysis was then strengthened with evidence collected through telephone interviews with 6 selected UNCTs and from the questionnaire completed in mid 2006 by 65 of the possible 80 UNCTs on the effectiveness of the guidelines. The evidence and findings were then triangulated against that found in other reviews of UNDAF performance.

Major findings and conclusions

Overall findings and conclusions

5. The present Guidelines establish a rigorous set of normative standards for monitoring and evaluation, although possibly due to constraints on the length of document allowed they do not provide detailed guidance on how to implement these normative standards. This finding also applies to the draft revision of the Guidelines circulated in November 2006 to a lesser extent.

6. The overall conclusion is that if fully implemented against these normative standards, the M&E approach proposed would significantly enhance the evaluability of the UNDAFs. This conclusion particularly applies to the definition and relevance of outcomes identified and the existence of sufficient measurable indicators. Norms for joint monitoring systems and especially for assessing risks and assumptions are comparatively less developed, although the draft revision of the Guidelines circulated in November 2006 does strengthen the treatment of key assumptions and risks.

7. However, it is important to understand that the normative standards implied by the Guidelines are extremely challenging to implement. In general, all of the UNDAF documents reviewed show that UNCTs (UN Country Teams) have attempted to follow the M&E approach suggested in the Guidelines, but the evidence and findings suggest that no UNDAF will fully meet the normative standards suggested in the Guidelines. Evaluation of the UNDAFs against the scope suggested in the Guidelines would require significant additional investment in addressing issues not adequately dealt with to date during the design and initial implementation of the UNDAFs. This conclusion is strengthened by the finding of MacKenzie (2006) that most UNCTs are not managing to implement their UNDAF monitoring plans, due to competing demands upon their time. However, it is also concluded that significant opportunities remain to further enhance the evaluability of the UNDAFs.

Clarity on the evaluation's purpose?

8. The UNDAF documents, Guidelines and UNCTs interviewed identify no clear and specific accountability demand that the evaluation should meet. Within the Guidelines, the implication is that the UNDAF evaluation should inform the development of the subsequent UNDAF through identification of lessons learned. However, the evidence suggests that UNCTs have given insufficient thought about when the evaluation would need to be scheduled to meet this purpose. The danger is that by the time they do consider this issue properly, it will be too late to implement an effective evaluation.

9. Evaluability is also adversely affected by a lack of clarity on what it is a priority to evaluate. The UNDAF is supposed to be the common strategic framework for the operational activities of the UN system at the country level and provide a collective, coherent and integrated UN system response to national priorities and needs, yet the Results Matrices and indicators do not adequately reflect this.

The infeasibility of measuring impact

10. The evidence rather suggests that most UNCTs want an evaluation to identify the impact of the UN. But review of both the documentation and discussion with UNCTs strongly suggests that most in the UNCTs have failed to appreciate that this would probably be impossible in the 4th year of a 5 year programming cycle. Quite simply, not enough time would have elapsed for the contribution to be detectable in any shift in national level indicators. If this is the major area of interest for UNCTs, then enhancing both evaluability and the utility of monitoring systems needs to start from the question of what it is technically feasible to monitor and then to evaluate during programme implementation and then ensure that M&E systems focus the limited resources at the appropriate level.

Sufficient measurable indicators

11. In general, UNCTs appear to have had great difficulty identifying relevant indicators that actually track performance and that are assigned at the correct level in the logical hierarchy. For example, at Country Programme Outcome level, few of the indicators actually identify specific results which an agency could be expected to realise within the time frame of the UNDAF and be held accountable against. The review also found that there is not the evidence required across the 32 UNDAFs reviewed for an evaluator to easily understand the logic between UN interventions, and associated indicators, and expected impact upon the national goals as shown in the Results Frameworks.

12. The M&E matrices also either have not included any UNDAF Outcome indicators or have pitched them at too high a level in the results hierarchy. This means that there is a lack of indicators on the added value from the UNDAF process itself in delivering a more collective, coherent and integrated UN system response.

Quality of joint monitoring systems

13. Neither the CCA-UNDAF Guidelines nor the UNDAF documents discuss the issue of either the capacity or resource requirements required to implement an M&E plan which will deliver against the normative standards implied in the Guidelines. Nor is there evidence that UNCTs have adequately considered how the demands for monitoring and evaluation of the UNDAF should affect the monitoring and evaluation activities of the individual agencies. This is a major over-sight, since until these two issues are considered any M&E plan is *de facto* no more than a statement of intent.

14. Discussion with six UNCTs also suggests that the M&E plans have not driven monitoring of the UNDAFs by the UN agencies and supports the findings of Mackenzie (2006), as stated below:

“Current practice is to re-organize theme groups (TGs) around the agreed UNDAF outcomes. These UNDAF outcome groups are meant to use the M&E framework as an operational tool and report on a regular basis to the UNCT about progress. This seems to be happening only rarely. It appears that the pressure for agency reporting to HQs, who are responsible for reporting to executive boards and to other governance bodies takes precedence almost immediately. This is understandable. But programme staff in the field have to fulfil their agency monitoring and reporting tasks, as well as participate meaningfully in theme groups, do joint monitoring and report on overall progress for achievement of the UNDAF outcomes. There is duplication, particularly because each agency maintains its own separate monitoring framework in the CPAP or other project instrument. Overall, it seems very difficult to do both well. Experience from the field suggests that the result is:

- Outcome groups that meet rarely and almost never meet with all members;
- Responsibilities that are usually not put into performance appraisal instruments;
- Group members who are not rewarded for their UNDAF monitoring and reporting efforts;
- Outcome groups that quickly become paper entities; and most importantly
- No regular or coherent reporting to the UNCTs about overall progress towards UNDAF outcomes.”

External factors

15. External factors (risks and assumptions) impact in two ways on the evaluability of the UNDAF:

- First, those risks and assumptions directly tied to the efficient and effective delivery of interventions by the UN agencies.

- Second, those that relate to what it is assumed will be delivered by others, if the UNDAF Outcome is to be delivered. This is required if an attempt is going to be made to evaluate the UN's contribution.

16. Review of the UNDAF documents showed that eight out of 33 UNDAFs include no reference to identified risks and assumptions, but more importantly, in no case, does any UNDAF discuss how risks and assumptions will be monitored. This significantly diminishes the evaluability of the UNDAF's, since if an evaluation cannot directly assess the impact of UN interventions on national goals, it can assess whether or not the assumptions on why the interventions should have such impacts are still valid.

17. The second set of assumptions relate to what others are doing. The UNDAFs do enhance evaluability in that they normally identify who the other key partners in achievement of the UNDAF outcomes are supposed to be. However, since the monitoring systems are not designed to track whether or not partners actually do what is expected, evaluability is still adversely affected.

Recommendations

18. The below recommendations focus upon enhancing the evaluability of the UNDAFs and therefore only discuss monitoring in this context. It is also difficult to identify specific recommendations, since their relevance will depend upon whether or not the conclusions and recommendations of the High Level Panel on System Wide Coherence are adopted and implemented.

Recommendation 1: The relevant sections of the Guidelines need a further revision.

19. It is recommended that any future revision of the evaluation and monitoring related aspects of the Guidelines needs to be based upon conclusions to the following four questions:

- Who needs what evidence, when and for what purpose? This would require finding out why UNCTs don't adequately address these issues when developing their UNDAF monitoring and evaluation plans.
- Is it methodologically possible to collect the needed evidence when it is needed and at what cost? The key recommendation in this report is that both monitoring and evaluation activities need to focus more on whether or not assumptions remain true, since timing means that evaluations or reviews will be formative in nature.
- Are the resources and capacity likely to be in place to derive the required evidence? Any plan that does not address these two issues during development is not a plan but a statement of intent and highly unlikely to be implemented, as is clearly seen with the present UNDAF M&E plans.
- What is the minimum level of monitoring and review activity that should be expected of an UNCT?

Recommendation 2: Drop the requirement for UNCTs to commission UNDAF evaluations.

20. Most UNCTs have neither the expertise nor the resources to commission a good quality evaluation (against the UNEG norms and standards). It is therefore recommended that UNCTs instead carry out less resource intensive reviews of performance against the UNDAF at the beginning of penultimate year as part of the process for developing a new UNDAF.

Recommendation 3: Establish what the likely demand for evaluative evidence actually is.

21. Assuming that UNCTs, in future, carry out a review of UNDAF implementation in the penultimate year, this study has not been able to identify any other current demand for evaluation of the UNDAF. How the demand for UNDAF evaluations for accountability purposes within the UN will develop is unknown and will depend upon whether the conclusions and recommendations of the High Level Panel on System Wide Coherence are adopted and implemented.

22. However, as a first step, and to improve demand for both monitoring and evaluation evidence, any future revision of the Guidelines should high-light the need for UNCTs to embed monitoring and evaluation of the UNDAF within the systems that the partner governments are creating, as proposed in the Paris Declaration, to monitor and evaluate donor and government performance against aid-effectiveness commitments.

Recommendation 4: Task UN evaluation functions to carry out UNDAF evaluations.

23. Given capacity constraints within both partner governments and UNCTs and the methodological challenges involved, UNDAF evaluations should be managed by UN evaluation functions. If such evaluations are to be commissioned, the central evaluation function should consider:

- The use of evaluability studies at the start of the UNDAF cycle in targeted countries as an approach to ensuring that sufficient evidence is collected and address the problems UNCTs have in implementing M&E plans.
- That no comparable evaluations have been carried out to date but that EvalNet's pending evaluation of harmonisation and alignment may provide concrete lessons on evaluation approaches and methodologies that would be relevant to any future UNDAF evaluation.

Recommendation 5: UNEG should review M&E guidance for the individual agencies and identify whether they can increase the availability of evidence that could underpin future UNDAF evaluations.

24. In practice, data and evidence for evaluation of an UNDAF is more likely to come from evaluations carried out by the individual UN agencies and by other stakeholders. Therefore, UNEG

should consider whether amending evaluation guidance issued by the individual UN agencies might not be the easiest way to enhance UNDAF evaluability.

Recommendation 6: Establish the principle that monitoring from the level of the CP Outcome downwards should be the responsibility of the individual implementing agencies and of national goals should lie with the partner government.

Recommendation 7: Focus the UNDAF monitoring framework and the M&E plan on what the added value of a coherent and coordinated UN country programme would be to delivery of national goals.

25. Assuming that the UNDAF actually does lead the programming processes within the individual UN agencies, the principle that monitoring from the level of the CP Outcome downwards should be the responsibility of the individual implementing agencies needs to be established more clearly. This would help ensure that scarce resources within the UNCT are not wasted on developing duplicate UNDAF monitoring systems at these levels and start to address the lack of attention to developing joint monitoring systems within the current Guidelines.

26. The focus of monitoring under the UNDAF process should then be to monitor the degree to which the overall UN programme is likely to add value:

- Through a more coherent and coordinated approach within UN country programme to delivery of national goals; and
- Through its influencing and political brokerage role in delivering the UNDAF Outcomes more effectively.

1. Background and objective of the study

27. In response to the Secretary-General's 1997 call for the United Nations to articulate a coherent vision and strategy for a unified approach towards common development goals at country level, the Common Country Assessment (CCA) and the United Nations Development Assistance Framework (UNDAF) were adopted as strategic planning tools for the UN system.

28. The UNDAF is the common strategic framework for the operational activities of the UN system at the country level. It aims to provide a collective, coherent and integrated UN system response to national priorities and needs, including PRSPs and equivalent national strategies, within the framework of the MDGs and the commitments, goals and targets of the Millennium Declaration, summits and instruments of the UN system. The UNDAF emerges from the analysis of the Common Country Assessment – an instrument used to analyze the national development situation and identify key development issues – and provides a basis for the preparation of UN system country programmes¹.

29. Initially (up to 2004), monitoring and evaluation (M&E) were not addressed in the guidelines issued by the UN Development Group on the development and use of UNDAFs. However, in 2002, an M&E task force was created that produced M&E recommendations in 2004, which were then incorporated into the guidelines issued. These revised guidelines should therefore have affected the 16 UNDAF documents finalised in 2004 and the 19 finalised in 2005.

30. This study is expected to provide inputs into the updating of the UNDAF M&E guidance, currently being prepared by UNDG led Working Group on Programming Policy. It is also meant to yield benefits for the on-going process of conducting joint country-level evaluations and to provide inputs to refine the scope of the next round of UNDAF evaluations that will be commissioned by the relevant UNCTs.

2. Scope and methodology

2.1 Scope

31. The study focuses on the 35 UNDAF documents (listed at Annex 1) completed in 2004 and 2005, where it can be expected that the revised Guidelines would have affected the approach adopted both to programme design and monitoring and evaluation and hence their evaluability.

2.2 Methodology

32. The ToRs for this study indicate that the study should “assess the evaluability of UNDAF in terms of: a) clarity of intent of the subject to be evaluated (relevance and design of the expected outcome statements and results matrices); b) the existence of sufficient measurable indicators (collection of reliable

¹U.N, Common Country Assessment and United Nations Development Assistance Framework – Guidelines for UN Country Teams preparing a CCA and UNDAF, undated, pp.11, 17.

data for analysis); c) quality of joint monitoring systems; and d) external factors (positive or negative) that have influenced the process and the realization of expected outcomes”.

33. Assessing evaluability requires making a judgement based on the evidence and therefore needs to be based on the use of commonly agreed normative standards wherever possible and the transparent application of said standards. Therefore, under the four broad questions identified in the ToRs, the UNEG Norms and Standards were used to develop a set of more detailed questions². Selection of relevant UNEG Standards was also influenced by review of the:

- Approach of UNDP when assessing the evaluability of the Assessment of Development Results for Serbia and Montenegro³.
- The paper outlining ‘Issues & Recommendations for the revision of the 2004 CCA-UNDAF Guidelines’ by A. MacKenzie (Sept. 2006).
- The ‘Review of the Role and Quality of the United Nations Development Assistance Frameworks (UNDAFs)’ by R. Longhurst (May 2006).
- Approach to assessing evaluability used by the IADB when assessing the evaluability of its country strategy documents⁴;

34. Selection of these questions also revealed that it would not be possible to interpret the evidence based on the purely technical aspects of the results framework, without making assumptions on what the purpose of the evaluation would be. Therefore a fifth group of questions was included that focused on the purpose of the evaluation. A listing of the questions used is at Annex 2.

35. The questions developed were then used to assess whether:

- i. The ‘Guidelines for UN Country Teams’ (2004) contain the elements required to allow the evaluation of the UNDAFs.
- ii. For the 35 UNDAF documents finalised in 2004 and 2005, to what degree these documents include the elements that would allow the future evaluation of these UNDAFs.
- iii. These documents actually followed the relevant suggestions in the ‘Guidelines for UN Country Teams’.

² The UNEG Norms only briefly touch upon evaluability, whilst evaluability is not mentioned in the UNEG Standards at all. However, many of the Standards directly deal with the issue of evaluability and therefore can be grouped under the four broad questions identified in the ToRs.

³ Assessment of Development Results – Serbia and Montenegro. Inception Report, August 2005.

⁴ Report on the Evaluability of Bank Country Strategies, Office of Evaluation and Oversight, Inter-American Development Bank, September 2005.

36. The above assessment was based upon a textual analysis of documentation. However, this approach has three key limitations. First, a lack of documentary evidence does not mean that something has not been done in reality. Second, the degree to which something can be evaluated depends upon the capabilities of those commissioning and implementing the evaluation and the level of resourcing that one is willing to commit to the evaluation. Third, the CCA and UNDAF documents are not the only sources of information and evidence that affect the evaluability of the UNDAF. Therefore, evidence on these three key issues was collected through telephone interviews with 6 selected UNCTs⁵ and from the questionnaire completed in mid 2006 by 65 of the possible 80 UNCTs on the effectiveness of the guidelines. The evidence and findings were then triangulated against that found in other reviews of UNDAF performance.

3. Findings

3.1 Do the ‘Guidelines for UN Country Teams’ (2004) contain the elements required to allow the evaluation of the UNDAFs?⁶

Finding 1: Do the Guidelines cover the purpose, scope and management of the evaluation?

37. The Guidelines are not clear on what the purpose of the evaluation is – purpose is defined here as being for either accountability or lesson learning or a mixture of the two. The suggested timing at the beginning of the penultimate year implies that the evaluation is intended to inform development of next UNDAF, which would suggest that its primary purpose is to learn lessons. However, Section 4.3 of the Guidelines, which covers the suggested scope of the evaluation, makes little reference to drawing lessons for the development of the next UNDAF. Neither does the guidance on the development of the CCAs make any substantive reference to using lessons learned from evaluations of past performance as a criterion for selection of future areas of cooperation.

38. The suggested scope of the evaluation is clearly set out in the Guidelines (Section 4.3), as shown below:⁷

⁵ Bangladesh, Serbia, Swaziland, Guyana, Armenia, and China.

⁶ Evidence supporting the below findings is summarised at Annex 3.

⁷ The scope of the UNDAF evaluation defined in the November 2006 draft revision of the Guidelines is as follows: (i) Did the UNDAF make the best use of the UN’s comparative advantages in the country?; (ii) Did the UNDAF generate a coherent UNCT response to national priorities?; and (iii) Did the UNDAF help achieve the selected priorities in the national development framework? This represents a narrowing of the scope to focus on areas (a), (b) and (f) of the evaluation scope identified within the current Guidelines.

- a. *Impact* –To what extent has progress in attaining the UNDAF Outcomes impacted national development? How has the progress in attaining the UNDAF outcomes benefited the poorest and most vulnerable people?
- b. *Relevance* - Did UNDAF outcomes strategically position the UN within the development community, especially in pursuit of national MDGs? Are the outcomes still valid for the next UNDAF?
- c. *Sustainability* – Are positive changes in the development situation sustainable? To what degree have strategies and programmes under the UNDAF been institutionalized? Have complementarities, collaboration and/or synergies fostered by the UNDAF contributed to sustainability? How have national capacities at the levels of government, local community, NGOs and civil society been enhanced?
- d. *Effectiveness vis-à-vis UNDAF and CP outcomes* – What progress has been made towards the CP and UNDAF outcomes and CP outputs?
- e. *Efficiency* – Were results achieved at reasonably low or lowest cost? How did strategies result in a more efficient, simplified and harmonised UN?
- f. *Effectiveness of the UNDAF as a coordination framework* - Has the UNDAF contributed to more complementary and collaborative programming by agencies? Did the UNDAF make programming by agencies more strategic and synergistic? Has value been added by these synergies? Has effectiveness been enhanced?

Finding 2: What relevant guidance is given on constructing a plausible logic model connecting UN support with national development goals?

39. The Results and Monitoring Matrices should effectively summarise:
 - What the UN agencies intend to do;
 - Expected results at UNDAF Outcome, Country Programme Outcome and Country Programme Output level;
 - How, in broad terms these link to particular national goals and outcomes;
 - Who the agencies will be working with (both inter-UN agency and other partners) and the relative contributions being made by the partners; and
 - What the UN agencies consider the main risks and assumptions to be.
40. They should therefore provide an evaluator with a baseline of the situation at the start of an UNDAF’s implementation and the logic assumed to underpin proposed interventions.
41. For the logic linking UN interventions with national goals, the CCA guidance clearly high-lights that the better the analysis of cause and effect is, the more accurate will be the resulting development

assistance frameworks in identifying steps necessary for the achievement of national and international goals. For instance the guidance points out that a statement about the “inadequacy of certain sector policies” is not sufficient to guide development interventions; at the minimum the missing or unsupportive elements should be indicated. In fact, a significant part of the guidance on CCAs is on approaches to analyzing the situation and identifying the immediate and root causes. Hence, the assumption should be that the detailed cause and effect logic required by an evaluation of the UNDAF outcomes should be found in the CCA, rather than in the UNDAF, given that it would be difficult to include this level of detail in an UNDAF that should not exceed 15 pages in length.

42. The Guidelines also state that ‘the UNDAF emerges from the analyses of the CCA and is the next step in the preparation of United Nations system country programmes and projects of cooperation’⁸. Within the Guidelines, the first identified step in development of the UNDAF is selection among, and agreement with Government on, areas of cooperation that will be taken forward by the UN. It is recommended that this process be carried out in a Prioritisation Workshop. However, since a record of the discussion within the Prioritisation Workshop is not appended to the UNDAF, the danger is that an insufficiently detailed rationale for selection of the areas of cooperation will be included in the UNDAF. In this case, insufficient is defined as meaning not enough evidence to allow robust examination of the logic underpinning selection of areas of cooperation.

43. It is during the Prioritisation Workshop that the UNCT are supposed to also identify between three and five UNDAF Outcomes.⁹ However, whilst Guidelines clearly define what an UNDAF Outcome is, they do not outline how the UNCT is supposed to derive the UNDAF Outcomes from the areas of cooperation. The danger is that not spelling out this difference clearly would mean that the two are seen as being the same thing, when in fact an UNDAF Outcome should be a much more specific and concrete statement of what results the UN will commit to deliver.

44. Finally, those responsible for assessing the quality of UNDAF’s should assess the robustness of the logic presented, when answering Question 7 in the Quality Review Template, which specifically asks whether ‘the UNDAF convincingly presents a strategic & results driven framework for UN agency programming to address poverty challenges prioritized in the CCA’.

Finding 3: What does the results matrix seek to measure?

45. The guidelines state that ‘linkages between national goals or targets (as related to specific MDGs and/or other international commitments, goals and targets) and UNDAF and Country Programme Outcomes are elaborated within the Results Matrix, together with resource requirements. For each outcome, the Matrix illustrates how the UNDAF guides the design of agency-supported Country

⁸ Part 3.1, Definition of an UNDAF

⁹ An UNDAF Outcome is defined as ‘the specific results which the United Nations system expects to realize within the time frame of the UNDAF as its contribution towards the achievement of the national development priorities and goals in each area of cooperation. The United Nations system is collectively accountable for these outcomes, working in collaboration with the Government and other development partners, and must be able to demonstrate progress towards their achievement’.

Programmes and projects and parallel or joint programming (see the Guidance Note on Joint Programming – 19 December 2003). Conversely the Matrix also shows how the major outcomes of agency-supported Country Programmes and projects lead to the achievement of the shared UNDAF outcomes that exceed the sum of individually planned efforts’.

46. Before, examining the UNDAF guidelines further, it is necessary to first be clear on what the expected added value of having UNDAF results and monitoring frameworks might be for an evaluator. Looking at the counterfactual scenario, it can be assumed that:

- Monitoring and analysis at the national goal level would occur anyway and is unlikely to be affected by the presence or absence of UNDAF results and monitoring frameworks and there is little reason to believe that the existence of the UNDAF leads to additional UN support in this area over and above that which would be provided by the individual UN agencies anyway.
- Monitoring and evaluation by the individual UN agencies would also occur anyway and given the lack of resourcing would be the major source of information on UN agency performance even in the presence of an UNDAF results and monitoring framework. Monitoring and evaluation of joint or collaborative programming between two or more agencies might be affected by the presence, or not, of the UNDAF results and monitoring frameworks, but the Guidelines do not discuss this issue.
- Prioritisation, at either the level of UNDAF Outcomes or below would be an exogenous factor to the results and monitoring frameworks. This is because prioritisation is dealt with before the results and monitoring frameworks are designed and the frameworks contain no information on what the alternatives open to the UNCT were.

47. For an evaluator, the added value from the matrix would logically be in the degree to which it can show how the major outcomes of agency-supported Country Programmes and projects lead to the achievement of the shared UNDAF outcomes that exceed the sum of individually planned efforts. This is effectively the *raison d’etre* for the UNDAF process. However, within the Guidelines, there is no guidance upon how UNCTs might define UNDAF Outcomes that reflect this additionality, nor is the issue of additionality flagged in the sections dealing with analysis during the CCA process or in the criteria suggested for prioritising and selecting between possible areas of cooperation. The risk therefore is that the results and monitoring matrices may fail to incorporate additionality and therefore this issue is not monitored.

Finding 4: What type of indicators does the Guidance suggest be used?

48. The Guidelines give a clear set of normative standards on what indicators should be included, as stated below¹⁰:

¹⁰ It should be noted that those responsible for monitoring the quality of the UNDAF document are not asked to assess the degree to which the indicators included in the UNDAF monitoring matrix meet these standards.

‘The framework lists, for each UNDAF outcome and related CP outcomes, one or more quantitative and/or qualitative indicator(s) for monitoring progress, including baseline data and sources of verification as well as risks and assumptions. CP outputs are also listed and indicators and baselines added when available. Indicators, including baselines, should be disaggregated by gender and any additional characteristics that may be relevant to disparities between population groups in the country.’

49. The Guidelines suggest that the indicators be formulated, together with the relevant national institutions, in such a way that progress towards their achievement can be objectively verified, by direct measurement or other means. The Guidelines however provide no examples of what indicators that could be used at either UNDAF or Country Programme Output level might look like or how they might be developed.

Finding 5: What Guidance is given on the use or collection of other evaluative evidence?

50. The Guidelines briefly note the need to identify data gaps, but this is not then reflected in the issues covered in the Monitoring & Evaluation Programme Cycle Calendar, so increasing the danger that UNCT’s will not identify gaps in time to ensure that measures are put in place to fill the gaps.

51. The Guidelines do not explicitly discuss the role of evaluation data from evaluations carried out by individual agencies in evaluation of the UNDAF. Nor do the Guidelines flag that such evaluations are likely to be the main sources of evidence upon which the evaluation of the UNDAF would be based.

Finding 6: Treatment of risks and assumptions

52. The Guidelines do suggest that risks and assumptions be identified in the monitoring matrix. However, the Guidelines do not suggest that programmes monitor these on an on-going basis or identify risk mitigation strategies.

3.2 Do recent UNDAF documents include the elements that would allow the future evaluation of these UNDAFs?

Finding 7: Definition of purpose and scope in UNDAF documents

53. The Guidelines were not explicit on what the purpose of the evaluation would be, but were clear on suggesting what its scope might be (see para. 13). Table 1 below presents an analysis from 33 of the 35 UNDAF documents agreed in 2004 and 2005 of what evaluation purpose was specified and the scope of the evaluation¹¹.

¹¹ Only 33 of the 35 UNDAF documents completed in 2004 and 2005 were included, since the reviewer’s Spanish wasn’t good enough to credibly review the two documents that were drafted in Spanish – Peru and Guatemala.

Table 1: Number of UNDAF documents (n=33) specifying the evaluation purpose and scope of the evaluation.

Scope of evaluation	Evaluation purpose				
	Lesson learning	Accountability	Both	Not Specified	Total
Impact on national goals	10	0	0	7	17
Relevance of UNDAF outcomes	10	0	0	4	14
Sustainability	1	0	0	2	3
Effectiveness of implementation	8	0	0	3	11
Efficiency of implementation	3	0	0	2	5
Effectiveness of the UNDAF as a coordination framework	7	0	0	4	11
Not specified	4	0	0	7	11

54. The table reveals four major findings:

- None of the UNCTs appear to see the function of the evaluation as being to meet accountability needs. This finding was confirmed in discussions with UNCTs, who did not identify any current specific and operational accountability function for the evaluation.
- Two UNCTs included no scheduled UNDAF evaluation within their UNDAF documents.
- Half of the UNDAF documents identified evaluating the impact of the UNDAF Outcomes on the national goals as a role of the evaluation.
- Approximately a third of the UNDAF documents identified evaluating the effectiveness of the UNDAF coordination framework as a role of the evaluation.

55. No UNDAF document identified in anything but the broadest terms who would be responsible for the planning and implementation of the evaluation. Nor were the milestones for the planning and implementation of the UNDAF evaluation included in the ‘Monitoring & Evaluation Programme Cycle Calendar’ included in finalised UNDAFs, as had been suggested in the Guidelines.

Finding 8: Do UNDAF documents include an evidence based rationale for selection of the UNDAF outcomes?

56. All UNDAF documents state that the UNDAF Outcomes were defined through Prioritisation Workshops and most state that the criteria identified in the Guidelines were used during the Prioritisation process. However, none of the documents provide details of what evidence was used and how these criteria were applied. The seven UNCTs contacted all stated that they had documented the proceedings of their Prioritisation Workshop and retained these in their records, which implies that the basic information required for evaluating the relevance of the Outcomes selected is available, even if not in the UNDAF document itself.

Finding 9: Is there sufficient evidence/information for an evaluator to develop or understand the logic between UN interventions and impact upon the national goals?

57. The main finding is that there is not the evidence required across the 32 UNDAFs reviewed for an evaluator to easily understand the logic between UN interventions and impact upon the national goals as shown in the Results Frameworks. Issues include that:

- In only 24 of the 32 UNDAFs were the problems impeding achievement of the national goals clearly identified. In the other eight cases the documents started from government goals/targets but did not identify the problems with achieving those targets/goals. However, in theory this issue can be dealt with by reference to both the CCA and PRSp documentation by an evaluator.
- In only 11 of the 32 cases was there any discussion of what the causes of the problem to be addressed were. In no case, was any evidence presented that would justify the assertions made of what the causes were, although review of a sample of CCAs suggests that it is possible in some cases to find detailed discussions of the causes, and sometimes supporting evidence.
- In only two cases, was there any discussion of why, and how, it was expected that the identified UN interventions would affect either the problem or cause identified. This makes development or review of the logic chain more challenging for the evaluator, because this information is not included in other documents either. In the case of the CCA, there is no requirement to discuss this issue. Programme or project documents, on the other hand, may look at the logic linking achievement of a particular CP Outcome with a national goal, but don't discuss the logic of why the sum of all of the CP Outcomes under one UNDAF Outcome will affect achievement of the national goal.

Finding 10: The utility for evaluation of UNDAF and CP Outcomes

58. As shown in Table 2, all UNDAF documents reviewed, barring one that included no results framework in the version publicly available, identify UNDAF and CP Outcomes. Thirty of the 32 frameworks also include CP Outputs.

Table 2: Number of UNDAF documents (n=32) including UNDAF Outcomes, CP Outcomes and CP Outputs by year of preparation

	2004 UNDAF documents	2005 UNDAF documents
UNDAF outcomes	14	18
CP outcomes	14	18
CP outputs	12	17
No results framework	1	-

59. However, the usefulness of the outcomes described for evaluation purposes is highly variable.

60. At the UNDAF Outcome level, UNCTs appear to find it is difficult to define Outcome statements that outline the specific results which the United Nations system expects to realize within the time frame of the UNDAF as its contribution towards the achievement of the national development priorities and goals in each area of cooperation. Nor do UNDAF Outcome statements identify the proposed additionality that is supposed to be derived from the interaction of results under the supporting CP Outcomes. Discussion with the seven UNCTs suggests that this may reflect the difficulty for UNCTs in defining precise UNDAF Outcome statements that are inclusive enough to cover the support across all UN agencies. This in turn reflects a tension under-lying the UNDAF approach. Namely, how do UNCTs balance the need to be inclusive and use the framework as an inclusive ‘big tent’ against using the UNDAF framework and process as a way of focusing and prioritising the combined efforts of the agencies. Finally, the utility of UNDAF Outcomes to an evaluator is severely compromised because nearly half of the UNDAFs include no indicators at this level, and, of the remainder, only 4 UNDAF’s include indicators that aren’t national government targets.

61. CP outcomes are intended to describe the intended results to which a specific agency-supported Country Programme, rather than the UN as a whole, contributes. CP Outcome statements are more specific and useful for evaluation purposes than those at UNDAF level.

Finding 11: The utility of UNDAF and CP Outcome indicators for tracking performance and delivery

62. Overall, monitoring matrices include examples of both quantitative and qualitative indicators. Matrices also either include baselines, where available, or note in the narrative that baselines will be established in the first year of implementation.

63. But as already discussed, in only 4 of the 33 UNDAFs, were UNDAF Outcome level indicators included that could be said to attempt to track performance by the UN¹². In other cases where indicators were included at this level, indicators were usually actually indicators of movement at the level of national goals, and therefore too high level to be useful as indicators at UNDAF Outcome level. Discussion with the seven UNCTs confirms the finding from MacKenzie (2006) that definition of indicators was difficult for UNCTs.

64. Whilst all UNDAF’s identified CP level outcome indicators, the major findings are that:

- Few of the indicators actually identified specific results which an agency could be expected to realise within the time frame of the UNDAF. In the 32 monitoring matrices reviewed, in only four cases did the CP outcome indicators consistently attempt to identify agency results, whilst in a further 11 cases this principle had been applied to some extent. In the remaining 17 cases, the judgement is that the indicators did not identify results at the required level.

¹²The criterion used here was that the indicators attempted to measure a change in institutional capacity.

- In only two cases, was there clear evidence that the matrix contained indicators that would allow some assessment of progress towards achievement of the CP Outcome whilst instances of indicators containing dates by when they should be achieved were rare.
- No indicators were included that monitored either whether strategies were resulting in a more efficient, simplified and harmonised UN or the effectiveness of the *UNDAF as a coordination framework*.

Finding 12: Do the UNDAFs identify what the expected contributions of other partners will be?

65. The UNDAF documents do identify who the UNCT expect the major partners to be, although this principle is inconsistently applied. For instance, some documents identify who the other major donors would be, whilst others make no mention of who these donors might be. Again, some UNDAF documents identify what they expect each partner's role to be, in broad terms, whilst most merely identify who the partners will be. None of the UNDAF documents identify the relative levels of inputs by each partner or identify which non-UNCT partners are key for delivery of UN Outcomes.

Finding 13: What evidence is there of planning for the evaluation?

66. Beyond stating when the evaluation will occur, there is no evidence in the UNDAF documents that UNCT's have considered the requirements for evaluation of the UNDAF when setting up the M&E systems. This was also confirmed in discussion with the seven UNCTs.

Finding 14: Treatment of risks and assumptions (external factors)

67. According to the Guidelines, risks and assumptions should be dealt with in two places in the UNDAF document. First, a description of major *risks and assumptions* which may affect the achievement of UNDAF outcomes should be included in the narrative. Second, risks and assumptions for each UNDAF and country programme/project outcome should be included in the Monitoring and Evaluation Framework. Review of the UNDAF documents shows that eight out of 33 UNDAFs include no reference to identified risks and assumptions. In no case, does any UNDAF discuss how risks and assumptions will be monitored.

3.3 Do recent UNDAF documents follow the relevant suggestions in the 'Guidelines for UN Country Teams'?

68. Table 3 below summarises the degree to which UNCTs have attempted to incorporate the relevant sections of the Guidelines within their UNDAF documents. Issues of the quality of what has been included are not addressed in the table, since it is covered in the section 3.2 above.

Table 2: Analysis of implementation of M&E guidance by year of UNDAF preparation

Included in UNDAF document?	2004 UNDAF		2005 UNDAF		Overall	
	Yes	No	Yes	No	Yes	No
1. Frameworks						
1.1. Results Matrix included	14	1	18	0	32	1
1.2 Result matrix includes all suggested topics	14	1	18	0	32	1
1.3 M&E Framework	11	4	17	1	28	5
1.4. M&E Framework follows suggested structure	9	6	15	3	24	9
1.5 M&E Programme cycle calendar	5	10	9	9	14	19
1.6 Inclusion of rationale for selection of UNDAF outcomes	15	0	18	0	33	0
2. Outcomes						
2.1 UNDAF outcomes	14	1	18	0	32	1
2.2 CP outcomes	14	1	18	0	32	1
3. Indicators						
3.1 UNDAF outcome	6	9	11	7	17	16
3.2 CP outcome	11	4	16	2	27	6
3.3 Baselines	7	8	7	11	14	19
3.4 Disaggregated by gender and vulnerable group	8	7	11	7	19	14
4. M&E Management						
4.1 Monitoring management identified	13	2	16	2	29	4
4.2. Building national monitoring capacity	14	1	18	0	32	1
4.3. Evaluation Management identified	9	6	12	6	21	12
4.4 Evaluation planning and implementation milestones	0	15	0	18	0	33
4.5 Scheduled for penultimate year	2	13	7	11	9	24

Finding 15: Do the UNDAF documents show evidence of being guided by the CCA-UNDAF Guidance?

69. In general, all of the UNDAF documents follow the structure suggested in the Guidelines and there appears to be little difference between those completed in 2004 and in 2005. With specific reference to monitoring and evaluation, the UNDAF documents also follow the Guidelines in the following areas:

- Structure and broad content of the Results and M&E Frameworks;

- Inclusion of UNDAF Outcome and CP Outcome statements;
- Inclusion of CP Outcome indicators;
- Inclusion of a rationale for choice of UNDAF outcomes;
- Definition of who will be responsible for monitoring during implementation of the UNDAF; and
- Building of national capacity.

70. UNDAF documents diverge from the Guidelines in the following areas:

- Inclusion of an M&E Programme Cycle Calendar;
- Inclusion of UNDAF Outcome indicators;
- Inclusion of baselines for indicators;
- Definitions of the evaluation planning and implementation milestones; and
- The timing of the evaluation.

3.4 Capacity and resourcing for evaluation of UNDAFs

Finding 16: The lack of evidence that capacity and resourcing for evaluation is considered during design of the UNDAF

71. Neither the CCA-UNDAF Guidelines nor the UNDAF documents discuss the issue of either the capacity or resource requirements required to implement an M&E plan which will deliver against the normative standards implied in the Guidelines. This is a major over-sight, since until these two issues are considered any M&E plan is *de facto* no more than a statement of intent. Discussion with the seven UNCTs suggests that the M&E plans have not driven monitoring of the UNDAFs by the UN agencies and support the findings of Mackenzie (2006), as stated below:

“ Current practice is to re-organize theme groups (TGs) around the agreed UNDAF outcomes. These UNDAF outcome groups are meant to use the M&E framework as an operational tool and report on a regular basis to the UNCT about progress¹³. This seems to be happening only rarely. It appears that the pressure for agency reporting to HQs, who are responsible for reporting to executive boards and to other governance bodies takes precedence almost immediately. This is understandable. But programme staff in the field have to fulfil their agency monitoring and reporting tasks, as well as participate

¹³ Please note that this is current practice – the guidelines do not describe or suggest this. However, the following are often referred to: [GlobalNet Digest discussion on Theme groups](#), September 2, 2003 to May 24, 2004, and [Basic principles for CCA/UNDAF theme groups](#), UNDG #5834, www.undg.org

meaningfully in theme groups, do joint monitoring and report on overall progress for achievement of the UNDAF outcomes. There is duplication, particularly because each agency maintains its own separate monitoring framework in the CPAP or other project instrument. Overall, it seems very difficult to do both well. Experience from the field suggests that the result is:

- Outcome groups that meet rarely and almost never meet with all members;
- Responsibilities that are usually not put into performance appraisal instruments;
- Group members who are not rewarded for their UNDAF monitoring and reporting efforts;
- Outcome groups that quickly become paper entities; and most importantly
- No regular or coherent reporting to the UNCTs about overall progress towards UNDAF outcomes”.

4. Conclusions and discussions

4.1 Overall conclusion

72. The ToRs for this study indicate that the study should “assess the evaluability of the UNDAF in terms of a) clarity of intent of the subject to be evaluated (relevance and design of the expected outcome statements and results matrices); b) the existence of sufficient measurable indicators (collection of reliable data for analysis); c) quality of joint monitoring systems and d) external factors (positive or negative) that have influenced the process and the realization of expected outcomes”.

73. The present Guidelines establish a rigorous set of normative standards for monitoring and evaluation, although possibly due to constraints on the length of document allowed they do not provide detailed guidance on how to implement these normative standards. This conclusion also applies to the draft revision of the Guidelines circulated in November 2006.

74. The overall conclusion is that if fully implemented against these normative standards, the M&E approach proposed would significantly enhance the evaluability of the UNDAFs. This conclusion particularly applies to the definition and relevance of outcomes identified and the existence of sufficient measurable indicators. Norms for joint monitoring systems and especially for assessing risks and assumptions are comparatively less developed. The draft revision of the Guidelines circulated in November 2006 is much more specific about the need to identify the key assumptions and risks, but it still only briefly mentions the need to develop risk mitigation strategies and does not identify the need to monitor the status of risks and assumptions on an on-going basis.

75. The results frameworks do describe intent at the beginning of the programming cycle; the CCA analyses should provide some of the under-pinning evidence and analysis required to create a more evidence based logic model against which to assess performance; and UNCTs have generally identified indicators and MoVs for tracking anticipated changes at the level of national goals. However, it is

important to understand that the normative standards implied by the Guidelines are extremely challenging to implement and the evidence and findings suggest that no UNDAF will meet the normative standards suggested in the Guidelines. It is therefore concluded that evaluation of the UNDAFs against the scope suggested would require significant additional investment in addressing issues not adequately dealt with during the design and initial implementation of the UNDAFs.

76. This conclusion is strengthened by the finding of MacKenzie (2006)¹⁴ that most UNCTs are not managing to implement the UNDAF monitoring plans, due to competing demands upon their time. Consequentially, many of the problems with the results frameworks will not become apparent until an attempt is made to evaluate against them. However, it is also concluded that significant opportunities remain to further enhance the evaluability of the UNDAFs.

4.2 Clarity on the evaluation's purpose?

77. The first evaluability issue deals with the ambiguity over what the purpose – accountability or lesson learning - of the UNDAF evaluation should be. Until those who have the resources to commission the evaluation and use the results are clear on why they need it, the probability is at best that the evaluation will not be cost-effective. Neither the UNDAF documents nor the Guidelines identify any clear and specific accountability demand that the evaluation should meet. Discussion with the seven UNCTs also confirmed that there is not a concrete demand from the partner governments for evaluation of the UNDAF and that there is no current demand for an evaluation to meet any internal UN accountability function¹⁵.

78. Within the Guidelines, the implication is that the UNDAF evaluation should inform the development of the subsequent UNDAF through identification of lessons learned¹⁶. However, of the 33 UNDAFs developed in 2004 and 2005, only nine had scheduled the evaluation in the penultimate year of the programme cycle, when it would be feasible to influence future UNDAF programming. By contrast, 17 UNDAFs had scheduled evaluations in the last year of the programme cycle¹⁷, two had scheduled no evaluation at all, whilst five had scheduled an evaluation in the middle programme cycle. The implication therefore is that UNCTs have given insufficient thought to when the evaluation would need to be scheduled, if it is to have any value. The danger is that by the time they do consider this issue properly, it will be too late to implement an effective evaluation.

79. The second issue impacting upon evaluability is the question of what precisely the evaluation should focus upon. The UNDAF is supposed to be the common strategic framework for the operational

¹⁴ MacKenzie, A. (2006). Issues and recommendations for the revision of the 2004 CCA-UNDAF Guidelines. A paper presented to the Working Group on Programming Policy. 26 September.

¹⁵ However, in future, it is expected that delivery against the UNDAF will become one aspect of RC performance that will be assessed. Whether this will increase demand for evaluative evidence is open to question.

¹⁶ This is also the assumption within the November 2006 draft revision of the Guidelines.

¹⁷ One UNCT suggested that scheduling the evaluation in the last year allowed the influencing of country programming by the individual agencies.

activities of the UN system at the country level and provide a collective, coherent and integrated UN system response to national priorities and needs. Yet, the findings suggest that evaluability of this issue has not been adequately addressed in the UNDAFs drafted in 2004 and 2005. This reflects:

- i. The reality that most UNDAF Outcome statements are too vague and do not focus on identifying what the added value might be from a more collective, coherent and integrated UN system response to addressing root causes of the development challenges; and
- ii. A failure to identify appropriate UNDAF outcome indicators in the monitoring matrices that measure either the added value from a more collective, coherent and integrated UN system response or from the UNCT's advocacy work.

80. This probably reflects the fact that UNCTs are not advised during the CCA/UNDAF process to look at the following two fundamental questions¹⁸:

- Why, and how, is it expected that identified UN interventions will affect either the problem or immediate/root cause identified?
- Where would be the added value be from a more collective, coherent and integrated set of UN interventions at impacting on the problems and its immediate/root causes?

4.3 The infeasibility of measuring impact

81. The third issue deals with the evaluability of the impact of UN interventions on national goals, which is the major evaluation issue identified in most UNDAFs for evaluation. The reality is that it is highly unlikely that any evaluation would be able to identify a contribution linking UN agency interventions identified in the UNDAF with shifts in the national goal level indicators, as is assumed in results frameworks, in the 4th year of a 5 year programming cycle. This is because if capacity building is the main focus of UN agency interventions, it is highly unlikely that the benefits will be discernible within such a short period of time. If this is the major area of interest for UNCTs, then enhancing evaluability, and the effectiveness of monitoring systems, needs to start from the question of what it is technically feasible to monitor and then to evaluate during programme implementation and then ensure that M&E systems focus the limited resources at the appropriate level. Review of both the documentation and discussion with UNCTs both strongly suggest that most in the UNCTs have failed to appreciate this limitation on what can and cannot be measured.

4.4 Sufficient measurable indicators

82. Conclusions on the adequacy of indicators fall into two classes. First, those related to the difficulty UNCTs appear to have encountered in identifying relevant indicators that actually track performance and that are assigned at the correct level in the logical hierarchy. Second, those related to a

¹⁸ The November 2006 draft revision of the Guidelines does.

lack of indicators on the added value from the UNDAF process itself in delivering a more collective, coherent and integrated UN system response.

83. UNCTs either have not included any UNDAF Outcome indicators or have pitched them at too high a level in the results hierarchy. This strongly suggests that UNCTs do not know how to identify such indicators, which is hardly surprising given that the need to define such indicators is a recent phenomenon. For example, UNCTs cannot fall back onto using M&E or RBM guidance from individual agencies', since these internal guidance documents also do not cover this area.

84. At CP Outcome level, few of the indicators actually identified specific results which an agency could be expected to realise within the time frame of the UNDAF. Drawing on the findings of MacKenzie (2006), the major conclusion is that this is a strong indicator of the disconnect between development of the UNDAF Results Framework, and its CP Outcomes, and the programming/monitoring processes within the individual UN agencies. However, it is likely that the solution to this problem will be through enhancing the RBM/M&E capacities within the individual agencies, rather than through the UNDAF process, given that it is at the level of the CP Outcome that the UNDAF M&E systems should connect with those of the individual UN agencies.

85. The lack of indicators on the added value from the UNDAF itself reflects a fundamental ambiguity within the monitoring and evaluation approach for the UNDAF itself. As stated earlier, the , the added value from the matrix would logically be in the degree to which it can show how the major outcomes of agency-supported Country Programmes and projects lead to the achievement of the shared UNDAF outcomes that exceed the sum of individually planned efforts. If UNCTs do not appreciate this issue, this may be a key reason why the perception has arisen that the UNDAF is creating parallel monitoring and evaluation tools and lines of reporting that are overwhelming UN capacities at country level, since the present monitoring UNDAF monitoring systems clearly don't focus on tracking and assessing the added value.

4.5 Quality of joint monitoring systems

86. The major finding was that beyond stating when the evaluation would occur, there is no evidence in the UNDAF documents that UNCT's have considered the requirements for evaluation of the UNDAF when setting up the UNDAF M&E systems. The conclusion is that this significantly lessens the evaluability of the UNDAFs, since the probability is that UNCTs will not identify the key challenges in evaluating the UNDAF in time to actually solve them.

87. This finding and conclusion reflects a probable wider lack of attention to the practicalities of managing the monitoring and evaluation of the UNDAF. Issues include:

- The lack of clarity identified above over what the added value of the results matrix is in terms of managing the UN response in a particular country;
- That neither the CCA-UNDAF Guidelines nor the UNDAF documents discuss either the capacity or the resource requirements required to implement an M&E plan which will deliver against the normative standards implied in the Guidelines.

- The lack of discussion of how the UNDAF monitoring system should relate to those of the individual agencies. The Guidelines don't flag the need to be clear on what the relative responsibilities and roles should be of the monitoring systems from the individual agencies as compared with that of the UNDAF system and how the two should link in practical terms.
- No discussion of how the ToRs of evaluations by individual agencies might be adjusted to enhance their value as sources of information for the UNDAF evaluation

4.6 External factors

88. External factors (risks and assumptions) impact in two ways on the evaluability of the UNDAF:

- First, those risks and assumptions directly tied to the efficient and effective delivery of interventions by the UN agencies.
- Second, those that relate to what it is assumed will be delivered by others, if the UNDAF Outcome is to be delivered. This is required if an attempt is going to be made to evaluate the UN's contribution.

89. Review of the UNDAF documents showed that eight out of 33 UNDAFs include no reference to identified risks and assumptions, but more importantly, in no case, does any UNDAF discuss how risks and assumptions will be monitored. This significantly diminishes the evaluability of the UNDAF's, since if an evaluation cannot directly assess the impact of UN interventions on national goals, it can assess whether or not the assumptions on why the interventions should have such impacts are still valid.

90. 65. The second set of assumptions relate to what others are doing. The UNDAFs do enhance evaluability in that they normally identify who the other key partners in achievement of the UNDAF outcomes are supposed to be. However, since the monitoring systems are not designed to track whether or not partners actually do what is expected, evaluability is still adversely affected.

5. Recommendations

91. The below recommendations focus upon enhancing the evaluability of the UNDAFs and therefore only discuss monitoring in this context. It is also difficult to identify specific recommendations, since their relevance will depend upon whether or not the conclusions and recommendations of the High Level Panel on System Wide Coherence are adopted and implemented.

Recommendation 1: The relevant sections of the Guidelines need a further revision

92. The overall conclusion is that both the 2004 Guidelines and the November 2006 draft revision impose a set of M&E related normative standards and requirements that cannot be met by the UNCTs. Nor do the Guidelines clearly specify a minimum subset of standards that all UNCTs could meet, given their likely level of capacity and access to resources for monitoring or evaluation. It is therefore

recommended that any future revision of the Guidelines needs to be based upon conclusions to the following four questions:

- Who needs what evidence, when and for what purpose? This would require finding out why UNCTs don't adequately address these issues when developing their UNDAF monitoring and evaluation plans.
- Is it methodologically possible to collect the needed evidence when it is needed and at what cost? The key recommendation in this report is that both monitoring and evaluation activities need to focus more on whether or not assumptions remain true, since timing means that evaluations will be formative in nature.
- Are the resources and capacity likely to be in place to derive the required evidence? Any plan that does not address these two issues during development is not a plan but a statement of intent and highly unlikely to be implemented, as is clearly seen with the present UNDAF M&E plans.
- What is the minimum level of monitoring and review activity that should be expected of an UNCT?

Recommendation 2: Drop the requirement for UNCTs to commission UNDAF evaluations

93. Good practice would endorse the need for the UNCT and in-country partners to commission and manage implementation of an exercise that assesses implementation and ensures that lessons learned are reflected in future programming exercises. However, most UNCTs have neither the expertise nor the resources to commission a good quality evaluation (against the UNEG norms and standards) to fulfil this function. It is therefore recommended that the requirement for UNCTs to commission an evaluation be amended to a requirement that they carry out a less resource intensive review of performance against the UNDAF at the beginning of penultimate year, with the following scope:

- Will the UNCT deliver the CP Outputs identified at the beginning of the programming cycle? This should mainly draw on data available from individual agency reporting systems.
- Is there any evidence that CP Outcomes will be delivered on time? In most cases, this would likely be a review on whether the assumptions on what needs to happen still hold true or whether the assumptions are invalid.
- Were the anticipated synergies between the agencies that were supposed to contribute to achievement of UNDAF Outcomes delivered?
- Is there evidence confirming that the UNCT's diagnosis of its comparative advantage during development of the UNDAF was correct?
- Were the assumptions used for selection of the UNDAF Outcome areas during the Prioritisation Workshop correct and do they still hold true?

- Do the UNCT's assumptions on why the delivery of the UNDAF Outcomes will make a significant contribution to delivery of national goals still hold?

Recommendation 3: Establish what the likely demand for evaluative evidence actually is.

94. Assuming that UNCTs, in future, carry out a review of UNDAF implementation in the penultimate year, this review has not been able to identify any other current demand for evaluation of the UNDAF. How the demand for UNDAF evaluations for accountability purposes within the UN will develop is unknown and will depend upon whether the conclusions and recommendations of the High Level Panel on System Wide Coherence are adopted and implemented.

95. However, as a first step, and to improve demand for both monitoring and evaluation evidence, any future revision of the Guidelines should high-light the need for UNCTs to embed monitoring and evaluation of the UNDAF within the systems that the partner governments are creating, as proposed in the Paris Declaration, to monitor and evaluate donor and government performance against aid-effectiveness commitments.

Recommendation 4: Task UN evaluation functions to carry out UNDAF evaluations.

96. Given capacity constraints within both partner governments and UNCTs and the methodological challenges involved, UNDAF evaluations should be managed by UN evaluation functions. If such evaluations are to be commissioned, the central evaluation function should:

- Consider the use of evaluability studies at the start of the UNDAF cycle in targeted countries as an approach to ensuring that sufficient evidence is collected.
- Consider that no comparable evaluations have been carried out to date but that EvalNet's pending evaluation of harmonisation may provide concrete lessons on evaluation approaches and methodologies that would be relevant to any future UNDAF evaluation.

Recommendation 5: UNEG should review M&E guidance for the individual agencies and identify whether they can increase the availability of evidence that could underpin future UNDAF evaluations.

97. In practice, data and evidence for evaluation of an UNDAF is more likely to be available from evaluations carried out by the individual UN agencies and by other stakeholders. Therefore, UNEG should consider whether amending evaluation guidance issued by the individual UN agencies might not be the easiest way to enhance UNDAF evaluability.

Recommendation 6: Establish the principle that monitoring from the level of the CP Outcome downwards should be the responsibility of the individual implementing agencies and of national goals should lie with the partner government.

Recommendation 7: Focus the UNDAF monitoring framework and the M&E plan on what the added value of a coherent and coordinated UN country programme would be to delivery of national goals.

98. Assuming that the UNDAF actually does lead the programming processes within the individual UN agencies, the principle that monitoring from the level of the CP Outcome downwards should be the responsibility of the individual implementing agencies needs to be established more clearly. This would help ensure that scarce resources within the UNCT are not wasted on developing duplicate UNDAF monitoring systems at these levels and start to address the lack of attention to developing joint monitoring systems within the current Guidelines.

99. The focus of monitoring under the UNDAF process should be to monitor the degree to which the overall UN programme is likely to add value:

- Through a more coherent and coordinated approach within UN country programme to delivery of national goals; and
- Through its influencing and political brokerage role in delivering the UNDAF Outcomes more effectively.

100. The degree to which UNCTs could monitor these issues would depend upon whether, during the CCA/UNDAF development process, they actually do look at the following two fundamental questions:

- Why, and how, is it expected that identified UN interventions will affect either the problem or immediate/root cause identified?
- Where would be the added value be from a more collective, coherent and integrated set of UN interventions at impacting on the problems and its immediate/root causes?

101. At least initially, it would also require significant monitoring expertise to be available through the UNDG, which could then help UNCTs to translate their understanding into a set of predefined indicators.

Annex 1: UNDAF documents completed in 2004 and 2005

2004	Angola	Madagascar	Iran Philippines	Guatemala	Armenia Azerbaijan Bosnia and Herzegovina Kazakhstan Kyrgyzstan Macedonia Romania Serbia Tajikistan Turkmenistan Uzbekistan
2005	Burkina Faso Cape Verde Chad	Ghana Namibia Swaziland Uganda	Afghanistan Bangladesh Cambodia China Indonesia Viet Nam	Guyana Peru	Albania Georgia Turkey Ukraine

Annex 2: Listing of key questions for documentary review

I. Context

Does the UNDAF include:

1. A clear explanation of the purpose of the UNDAF evaluation?
2. A clear definition of the roles and responsibilities in the planning and implementation of the evaluation?
3. Who the main intended users are likely to be?
4. What the main information needs of the intended users of the evaluation are likely to be?
5. When in the UNDAF programme cycle the evaluation should be carried out?
6. The evaluation criteria that should be used by the evaluators?
7. Who the other major stakeholders are for delivery against the UNDAF, their expected roles and level of contribution.
8. What the scope (outlining what is covered and what is not covered by the evaluation) should be?

II. Clarity of intent of the subject to be evaluated (relevance and design of the expected outcome statements and results matrices)

Does the UNDAF include:

9. A clear explanation of how UNDAF Outcomes are derived from the diagnostic work carried out in the CCA and/or other processes?
10. Does the UNDAF narrative or results framework identify what the causes of the problems identified are and the evidence supporting these assumptions?
11. Does the UNDAF narrative or results framework identify why the identified interventions will impact positively upon the causes identified within the life of the UNDAF?
12. Does the UNDAF identify UNDAF outcomes at the appropriate levels? The UNDAF outcome will be formulated as an expected change in institutional capacity for achieving the national priority or goals, or as behavioural change. It may represent a contribution at either the national or sub-national level.
13. Does the UNDAF identify CP outcomes at the appropriate level? CP outcomes describe the intended results to which a specific agency-supported Country Programme contributes.

14. Does the UNDAF identify CP outputs at the appropriate level? These refer to specific products or services resulting from development interventions. They can normally be drafted when elaborating the Country Programme document.
15. Does the Results Framework specify CP outcomes for all of the identified UN agencies, together with the contributions of other partners, which would allow one to examine whether there is a reasonable chance that they will lead to the attainment of the UNDAF outcome.
16. Conclusion: Are the elements present in the results framework which would allow one to assess whether logically and plausibly related?

III. Existence of sufficient measurable indicators and collection of reliable data for analysis

17. Do the results and/or M&E matrices identify quantitative or qualitative measure(s) of the results expected?
18. Are baselines included for relevant indicators?
19. Are these disaggregated by gender and any additional characteristics that may be relevant to disparities between population groups in the country?
20. Do the UNDAF Outcome indicators identify the specific results which the United Nations system expects to realize within the time frame of the UNDAF as its contribution towards the achievement of the national development priorities and goals in each area of cooperation?
21. Do the CP Outcome indicators identify the specific results which the identified agency expects to realize within the time frame of the UNDAF as its contribution towards achievement of the UNDAF outcome?
22. Are there indicators included that measure progress towards the identified outcome as well as achievement of the outcome?
23. Are the indicators relevant for assessing performance?
24. Was the guidance followed in this UNDAF?

IV. Quality of joint monitoring systems

25. Does the UNDAF include Programme Cycle Calendar?
26. Does the UNDAF specify what data gathering systems should be in place to generate information on indicators?
27. Identify who will be responsible for collecting the necessary monitoring data?

28. Does the UNDAF state whether the resources needed to ensure that the predefined data will be collected and analyzed in a timely manner are in place?
29. Does the UNDAF specify what provisions will be put in place to help other partners strengthen their data collection and processing mechanisms from the onset of the UNDAF's execution so as to facilitate their own monitoring and evaluation of the UNDAF?
30. Identify data requirements that cannot be met through existing monitoring and evaluation systems and responsibility for meeting these needs?
31. Used the Guidelines?

V. External factors (positive or negative) that have influenced the process and the realization of expected outcomes.

32. Identify what the enabling conditions for successful execution of the UNDAF are?
33. Identify specific indicators to track the risks and assumptions?
34. Who will monitor whether these enabling conditions (assumptions and risks) change?

Annex 3: Evidence for assessing how evaluability is addressed in the 2004 CCA-UNDAF Guidelines

Question	How addressed in guidelines
I. Purpose of the Evaluation Does the guidance include:	
A clear explanation of the purpose of UNDAF evaluations?	<p>Partial. Statement below does state what broad scope and purpose are, who main parties are and when should be done. But only partial since not explicit about how the evaluation should be used. Timing at the beginning of the penultimate year implies should inform development of next UNDAF, but what identified as the evaluation focus in Section 4.3 of the Guidelines is backward looking and no emphasis on learning lessons for designing next UNDAF. Therefore is the purpose to strengthen accountability (and if so to whom) or lesson learning and which is the more important? Note that use of the word 'review' to describe the UNDAF evaluation causes some ambiguity, since within the Guidance a review is defined as something different from an evaluation, see Table 4.2 for example.</p> <p>The UNDAF evaluation is a joint UN review, conducted with national partners, of the overall results of the UNDAF programming cycle. The evaluation also assesses whether the UNDAF was effective as a tool to support achievement of national priorities and to enhance coordination and harmonisation among all UN agencies. The UNDAF evaluation takes place at the beginning of the penultimate year of the UNDAF cycle. (Section 4.2, part 2)</p>
A clear definition of the roles and responsibilities in the planning and implementation of the evaluation?	<p>Partial. See comment below.</p> <p>The UNCT may decide to establish an M & E working group, possibly including national and other partners, to oversee the monitoring and evaluation of the UNDAF. (Section 4.2, part 2)</p>
Who the main intended users are likely to be?	<p>Partial. Implicit is assumption that it will be the UNCT and hopefully national government. Table 4.2 includes row on 'Use of information' that might expand on this.</p>
What the main information needs of the intended users of the evaluation are likely to be?	<p>No.</p> <p>A range of evaluation questions are identified in Section 4.3 of the guidelines. Whilst all of the evaluation questions are valid, without being clear on the ultimate purpose of the evaluation it would be difficult to assess the relative priority between these.</p>

Question	How addressed in guidelines
When in the UNDAF programme cycle the evaluation should be carried out?	<p>Yes.</p> <p>At beginning of penultimate year, which would imply that just before CCA process.</p>
The evaluation criteria that should be used by the evaluators?	<p>Yes. See quote from Section 4.3 in guidelines below.</p> <p>The UNDAF final evaluation should respond to the following concerns:</p> <ul style="list-style-type: none"> a) <i>Impact</i> –To what extent has progress in attaining the UNDAF Outcomes impacted national development? How has the progress in attaining the UNDAF outcomes benefited the poorest and most vulnerable people? b) <i>Relevance</i> - Did UNDAF outcomes strategically position the UN within the development community, especially in pursuit of national MDGs? Are the outcomes still valid for the next UNDAF? c) <i>Sustainability</i> – Are positive changes in the development situation sustainable? To what degree have strategies and programmes under the UNDAF been institutionalized? Have complementarities, collaboration and/or synergies fostered by the UNDAF contributed to sustainability? How have national capacities at the levels of government, local community, NGOs and civil society been enhanced? d) <i>Effectiveness vis-à-vis UNDAF and CP outcomes</i> – What progress has been made towards the CP and UNDAF outcomes and CP outputs? e) <i>Efficiency</i> – Were results achieved at reasonably low or lowest cost? How did strategies result in a more efficient, simplified and harmonised UN? f) <i>Effectiveness of the UNDAF as a coordination framework</i> - Has the UNDAF contributed to more complementary and collaborative programming by agencies? Did the UNDAF make programming by agencies more strategic and synergistic? Has value been added by these synergies? Has effectiveness been enhanced?
Who the other major stakeholders are for delivery against the UNDAF, their expected roles and level of contribution.	<p>Yes. Should be specified in both the results section of the UNDAF and also in the UNDAF matrix. See below statement from Part 3.3, Section 2 of the Guidelines.</p> <p>An initial identification of the United Nations agencies, national and other partners that expect to contribute to the attainment of each UNDAF outcome and corresponding Country Programme and project outcomes and strategies, and the coordination and programme/project modalities that they propose to employ, including mobilization of resources.</p>

Question	How addressed in guidelines
<p>II. Clarity of intent of the subject to be evaluated (relevance and design of the expected outcome statements and results matrices)</p> <p>Does the guidance identify the need for the results framework to link the results framework with the diagnosis of the problems to be addressed, by:</p>	
<p>Explaining how UNDAF Outcomes are derived from the diagnostic work carried out in the CCA and/or other processes?</p>	<p>Partial</p> <p>Road Map on page 5 of Guidelines shows that CCA should culminate in identification of a shortlist of development challenges, which should then form basis for discussion during a Prioritisation Workshop between the UNCT and government at which consensus is reached on the top 3 to 5 priorities, which are then reflected in the UNDAF Outcomes. Prioritisation is guided by criteria, such as the magnitude and growth of the challenge; whether a national commitment exists; whether the challenge falls within the range of UN agencies' mandate and comparative advantage; whether the possibility for synergies between the efforts of partners exists; and, most importantly, where the UN - acting together - can make the biggest difference.</p> <p>Note issue here that Guidelines do not discuss what the practical difference between an area of cooperation as defined in the CCA and an UNDAF outcome is.</p>
<p>Clearly identifying what the problems to be addressed are?</p>	<p>Yes. This is a major focus of Part 2 of the Guidelines.</p> <p>The UNDAF is the common strategic framework for the operational activities of the United Nations system at the country level. It provides a collective, coherent and integrated United Nations system response to national priorities and needs, including PRSPs and equivalent national strategies, within the framework of the MDGs and the commitments, goals and targets of the Millennium Declaration and international conferences, summits, conventions and human rights instruments of the UN system. The UNDAF emerges from the analyses of the CCA and is the next step in the preparation of United Nations system country programmes and projects of cooperation. (Part 3.1, Definition of an UNDAF)</p>
<p>Clearly identifying what the causes of the problems identified are and the evidence supporting these assumptions?</p>	<p>Yes. Part 2.3 in Guidelines, Analysis of Root Causes does address this issue.</p> <p><i>The quality of the CCA and, hence, the relevance of subsequent country programming depends on the depth and quality of the analysis.</i> The greater the specificity in analysing cause and effect, the more accurate will be the resulting development assistance frameworks in identifying steps necessary for the achievement of national and international goals. For instance a statement about the "inadequacy of certain sector policies" is not sufficient to</p>

Question	How addressed in guidelines
	guide development interventions; at the minimum the missing or unresponsive elements should be indicated.
The logic (model) for why the identified interventions will impact positively upon the causes identified?	<p>Partially. It really depends upon (i) the degree to which the analysis carried out in the CCA process is carried through into the UNDAF framework and outcomes but more importantly into the contributing country programme outcomes for the individual agencies and (ii) how effective those responsible for assessing the quality of UNDAF's are at addressing question 7 in the UNDAF Review template.</p> <p>Does the UNDAF convincingly present a strategic & results driven framework for UN agency programming to address poverty challenges prioritized in the CCA? (Question 7, Guidelines Annex 4B)</p>
The need to define objectives which are clearly linked to the problems and causes identified in the diagnosis.	<p>Yes. See quote below</p> <p>The linkages between national goals or targets (as related to specific MDGs and/or other international commitments, goals and targets), and UNDAF and Country Programme Outcomes are elaborated within the Results Matrix, together with resource requirements. For each outcome, the Matrix illustrates how the UNDAF guides the design of agency-supported Country Programmes and projects and parallel or joint programming (see the Guidance Note on Joint Programming – 19 December 2003). Conversely the Matrix also shows how the major outcomes of agency-supported Country Programmes and projects lead to the achievement of the shared UNDAF outcomes that exceed the sum of individually planned efforts. (Part 3.2 under discussion of the Results Section of the UNDAF)</p>
How all components are expected to contribute to the achievement of the objectives.	Yes. See quote immediately above.
That all elements in the results framework are logically and plausibly related.	<p>Yes. Those responsible for assessing the quality of UNDAF's should look at this.</p> <p>Does the UNDAF convincingly present a strategic & results driven framework for UN agency programming to address poverty challenges prioritized in the CCA? (Question 7, Guidelines Annex 4B)</p>
Who the beneficiaries will be?	<p>Yes. See below statement from Part 3.3, Section 2 of the Guidelines.</p> <p>This is the core section of the UNDAF. It describes <i>what</i> the United Nations system expects to</p>

Question	How addressed in guidelines
	focus on; <i>why</i> it focuses on these areas; <i>how</i> the expected outcomes will be achieved and <i>with whom</i> . Each selected area of cooperation in the results framework addresses these four questions.
What the CT might do to improve the clarity of definition of outcomes and outputs in the results framework.	Yes. See Table 3.1. UNDAF Results Matrix (Format for each UNDAF outcome)
What the key evaluation questions might be	<p>Yes. See quote from Section 4.3 in guidelines below.</p> <p>The UNDAF final evaluation should respond to the following concerns:</p> <p><i>Impact</i> –To what extent has progress in attaining the UNDAF Outcomes impacted national development? How has the progress in attaining the UNDAF outcomes benefited the poorest and most vulnerable people?</p> <p><i>Relevance</i> - Did UNDAF outcomes strategically position the UN within the development community, especially in pursuit of national MDGs? Are the outcomes still valid for the next UNDAF?</p> <p><i>Sustainability</i> – Are positive changes in the development situation sustainable? To what degree have strategies and programmes under the UNDAF been institutionalized? Have complementarities, collaboration and/or synergies fostered by the UNDAF contributed to sustainability? How have national capacities at the levels of government, local community, NGOs and civil society been enhanced?</p> <p><i>Effectiveness vis-à-vis UNDAF and CP outcomes</i> – What progress has been made towards the CP and UNDAF outcomes and CP outputs?</p> <p><i>Efficiency</i> – Were results achieved at reasonably low or lowest cost? How did strategies result in a more efficient, simplified and harmonised UN?</p> <p><i>Effectiveness of the UNDAF as a coordination framework</i> - Has the UNDAF contributed to more complementary and collaborative programming by agencies? Did the UNDAF make programming by agencies more strategic and synergistic? Has value been added by these synergies? Has effectiveness been enhanced?</p>
What the scope (outlining what is covered and what is not covered by the evaluation) should be?	Yes. See immediately above.

Question	How addressed in guidelines
<p>III. Existence of sufficient measurable indicators and collection of reliable data for analysis</p> <p>Does the UNDAF guidance identify the following?</p>	
<p>That the indicators included in the results matrix identify relevant quantitative or qualitative measure(s) of the results expected.</p>	<p>Yes.</p> <p>See section 4.2 in guidelines:</p> <p>An M&E framework. This is a management tool that brings together key M & E information in one table for easy and consistent reference for the UNCT and partners. The framework lists, for <u>each UNDAF outcome</u> and related CP outcomes, one or more quantitative and/or qualitative indicator(s) for monitoring progress, including baseline data and sources of verification as well as risks and assumptions. CP outputs are also listed and indicators and baselines added when available. Indicators, including baselines, should be disaggregated by gender and any additional characteristics that may be relevant to disparities between population groups in the country. A suggested format, to be used for each UNDAF outcome, for the M&E framework follows in table 4.1</p> <p>And see section 6.3 in the Guidelines:</p> <p><i>UNDAF outcomes:</i> The strategic focus of country programmes and projects will be determined from among the expected outcomes in the UNDAF. These UNDAF outcomes are directly linked to a national priority or goal. The selected national goal or target should relate to specific MDGs and/or the other commitments, goals and targets of the Millennium Declaration, and international conferences, summits, conventions and human rights instruments of the UN system. The UNDAF outcomes will be reflected in the main areas of focus and priorities of the agency's country programmes and projects. There might be some areas in the agency's programmes and projects related to their specific mandates, including those related to their normative roles, which go beyond the scope of UNDAF outcomes.</p> <p><i>Country programme/project outcomes:</i> <i>The outcomes of individual agency's country programmes/projects describe the intended results, which contribute to the UNDAF outcomes. This will ensure that the results of country programmes and projects are linked with the expected outcomes in the UNDAF.</i></p> <p><i>Country programme/project outputs:</i> The outputs of country programmes/projects are the specific products and/or services for which UN agencies are accountable, and which</p>

Question	How addressed in guidelines
	contribute to the expected outcomes of the country programme/project, as well as to UNDAF outcomes. The outputs from different country programmes will be more complementary and together lead to the achievement of the UNDAF outcomes. It is also desirable that the ground level programmes and projects of each UN agency be meaningfully related to those of other UN agencies actively working in the same area. These inter-relationships should normally be reflected in the programme and project documents concerned.
That baselines are established for relevant indicators	Yes. See above.
That indicators should be defined that measure progress towards the identified outcome as well as achievement of the outcome.	Yes The framework lists, for <u>each UNDAF outcome</u> and related CP outcomes, one or more quantitative and/or qualitative indicator(s) for monitoring progress, including baseline data and sources of verification as well as risks and assumptions. CP outputs are also listed and indicators and baselines added when available. Indicators, including baselines, should be disaggregated by gender and any additional characteristics that may be relevant to disparities between population groups in the country.
Give a clear definition of what an outcome and an output mean in the context of the results framework?	Yes See Tables 3.1 and 4.1 in Guidelines.
IV. Quality of joint monitoring systems Does the UNDAF guidance:	
Specify what data gathering systems should be in place to generate information on indicators?	Yes. See 4.2 in guidelines The UNDAF M&E Plan consists of three elements: 1. An M&E narrative , in the UNDAF document, describing how the UNCT will undertake and coordinate <i>monitoring of the UNDAF</i> : <ul style="list-style-type: none"> • a description of <i>coordination mechanisms</i> (e.g. theme groups, joint field visits, and other review activities with partners), stating lines of responsibility and accountability for oversight and completion of M&E tasks 2. An M&E programme cycle calendar. This is an implementation tool to improve coordination of UN M&E activities, enhance interagency collaboration in M&E, identify gaps in data collection and highlight how and when products of UN M&E

Question	How addressed in guidelines
	<p>activities will be used. The M&E programme cycle calendar schedules all major M&E activities (surveys/studies, assessments, reviews, M&E capacity building) and articulates how and by whom outcome achievements will be measured, UNDAF evaluation milestones, uses and users of information, and complementary partner activities. A suggested format for the M&E programme cycle calendar is given in table 4.2.</p>
Identify who will be responsible for collecting the necessary monitoring data?	Yes. See above.
Specify that resources should be identified and committed to ensure that the predefined data will be collected and analyzed in a timely manner.	<p>No.</p> <p>Under section 4.2 does say that:</p> <p>Other considerations that guide M&E activities include a balance of costs, national capacity building efforts and the likely benefits.</p>
Specify how to assess what level of resources would be needed to collect the pre-defined date.	No.
Specify that adequate provisions need to be put in place to help other partners strengthen their data collection and processing mechanisms from the onset of the UNDAF's execution so as to facilitate their own monitoring and evaluation of the UNDAF?	<p>Yes. Under 4.2 in guidelines states that:</p> <p>The UNDAF M&E Plan consists of three elements:</p> <ol style="list-style-type: none"> 1. An M&E narrative, in the UNDAF document, describing how the UNCT will undertake and coordinate <i>monitoring of the UNDAF</i>: <ul style="list-style-type: none"> • a description of <i>coordination mechanisms</i> (e.g. theme groups, joint field visits, and other review activities with partners), stating lines of responsibility and accountability for oversight and completion of M&E tasks • a description of <i>efforts to strengthen national M&E capacities</i> including timing and partners involved. Areas identified in the CCA and MDGRs for strengthening national monitoring systems, including disaster-preparedness measures, should be included. • a description of major <i>risks and assumptions</i> which may affect the achievement of UNDAF outcomes
That data requirements that cannot be met through existing monitoring and evaluation systems are clearly identified and	No.

Question	How addressed in guidelines
responsibility for meeting these needs allocated to the appropriate partner?	
That the evaluation methodologies to be used for data collection, analysis and involvement of stakeholders should be appropriate to the subject to be evaluated, to ensure that the information collected is valid, reliable and sufficient to meet the evaluation objectives, and that the assessment is complete, fair and unbiased.	No.
V. External factors (positive or negative) that have influenced the process and the realization of expected outcomes. Does the UNDAF guidance:	
Identify what the enabling conditions for successful execution of the UNDAF are?	Yes. See section 4.2 in guidelines. The UNDAF M&E Plan consists of three elements: <ol style="list-style-type: none"> 1. An M&E narrative, in the UNDAF document, describing how the UNCT will undertake and coordinate <i>monitoring of the UNDAF</i>: <ul style="list-style-type: none"> • a description of <i>coordination mechanisms</i> (e.g. theme groups, joint field visits, and other review activities with partners), stating lines of responsibility and accountability for oversight and completion of M&E tasks • a description of <i>efforts to strengthen national M&E capacities</i> including timing and partners involved. Areas identified in the CCA and MDGRs for strengthening national monitoring systems, including disaster-preparedness measures, should be included. • a description of major <i>risks and assumptions</i> which may affect the achievement of UNDAF outcomes
Identify specific indicators to track the risks and assumptions?	No.
That the monitoring system should also monitor whether these enabling conditions (assumptions and risks) change?	No.
Assign explicit responsibility for monitoring these risks?	No.