

## **2007 survey of the DAC Evaluation Quality Standards (for test phase application)**

### **Background**

The DAC Evaluation Quality Standards have been developed by the members of the Evaluation Network.<sup>1</sup> The standards were approved in March 2006 for a test phase application of three years and they are now being used by a variety of evaluation stakeholders. Experience during the test phase will inform final agreement on the standards. (In this document the DAC Evaluation Quality Standards are referred to as the standards, or in places DAC standards to avoid confusion.)

The standards are a guide to good practice and identify the key pillars needed for a quality evaluation process and product. They aim to contribute to a harmonised approach to evaluation in line with the principles of the Paris Declaration on Aid Effectiveness.<sup>2</sup>

### **The survey**

#### **Objective and methods**

At the sixth meeting of the DAC Network on Development Evaluation on the 27-28 June 2007, it was agreed that the Secretariat should develop a survey with the objective to gather initial feedback on the standards and to obtain a better picture of current efforts in using them. The information gathered from the survey would feed into the process leading up to final agreement on the standards.

The online survey was conducted in the period of July and September 2007 and was based on a questionnaire on the use of the standards. The questionnaire also addressed the challenges and constraints experienced when using the standards. In addition, the questionnaire included an optional question under which specific comments on the content of the standards and feedback on particular items could be provided. The questionnaire was sent via e-mail to all the members of the DAC Network on Development Evaluation and it was suggested that it be forwarded on to non-members. The survey was filled in by 21 of the Network members (out of which two were multilaterals) and by four other organisations (Preparatory Commission for the Comprehensive Nuclear-Test-Ban Treaty Organisation, UNICEF, WFP and the Ministry of Science and Technology in China).

The key survey findings are summarised below and is followed by an analysis of the responses. Annex 1 contains the survey questions and a summary of the answers. Annex 2 provides feedback on the quality and usefulness of the standards.

#### **Main findings**

- Eighty per cent of the respondents stated that they are making use of the DAC Evaluation Quality Standards.
- Central evaluation offices and departments are the main user of the standards, but consultants also frequently use them.
- Many agencies specifically refer to the standards for quality control, when drafting terms of references, in preparing evaluation guidelines, for staff training on evaluation and in briefing consultants.

---

<sup>1</sup> The DAC Evaluation Quality Standards (for test phase application)  
<http://www.oecd.org/dataoecd/51/7/38686953.pdf>

<sup>2</sup> Paris Declaration on Aid Effectiveness: Ownership, Harmonisation, Alignment, Results and Mutual Accountability.  
<http://www.oecd.org/dataoecd/11/41/34428351.pdf>

- Members generally see the DAC standards as helpful in setting internationally agreed standards. Some respondents report that the standards have been useful in preparing internal guides and approaches.
- Although many believe that the standards have improved evaluations, a number of respondents also noted that the standards have not been used long enough to make a definite judgement on quality.
- The majority of the agencies are not recording experiences of the use of the standards as it is seen too early to do so, however, many stated that they will start record experiences in early 2008.

## Results

Twenty-one of the evaluation network members participated in the survey, indicating that there is great general awareness of, and interest in, the DAC Evaluation Quality Standards among members. Four of the respondents to the survey are not members of the evaluation network, indicating that awareness of the standards is somewhat spread outside the network.

The standards have been taken on board to various degrees by the organisations responding to the survey. Those organisations that have not entirely adopted the DAC standards either often draw upon the DAC standards while developing their own standards or state that the standards in use in their organisation are similar to the DAC standards. This implies that although the DAC standards might not be used in its totality, they could be used to build and improve organisations' own evaluation standards and that the DAC standards are in coherence with already existing best practice of various organisations.

Moreover, the survey reveals that the DAC standards are referred to at various stages of the evaluation process and in particular in drafting terms of references, quality control, writing evaluation guidelines, staff training on evaluation and when briefing consultants. Some agencies seem to refer to the standards more often and include the standards in all areas of work. Others state that the standards are taken, or will be taken, into account when drafting their own standards. These responses confirm that the DAC Evaluation Quality Standards are well known and used by many agencies.

The majority of the organisations responding to the survey reported that the central evaluation departments are the main users of the standards. The survey further indicates that consultants are another main group using the standards. Most agencies' evaluation departments who use the standards are also promoting the use of them to other departments within their organisation as well as to partners. However, it is important to note that even though the standards have been shared with country and program departments within a specific agency, it does not necessarily mean that the standards are being used within these other departments. The survey in fact shows that only a small number of the country and program departments use the standards. It is also indicated that the use of the standards is low for field offices and embassies. However, it is important to keep in mind that the respondents are mostly from evaluation departments and hence might not have a detailed knowledge of what is actually used in the field or by other departments undertaking evaluation work.

At this point only a few agencies are recording experiences of the DAC standards, and most of those who do so, are not doing so regularly. The main reason given for not recording experiences is that the standards have not been used for very long and that the agencies require more experience before they start recording them. Seven agencies explicitly stated that they will start recording experiences in early 2008, when the standards have been used for some time.

Many organisations are stating that the standards are useful and it is generally believed they will improve evaluations. The fact that the standards are internationally agreed upon makes it easier to convince consultants and others to abide by them. Some agencies state that the standards have already improved communication with, and work by, consultants. It is also stated that the standards have improved harmonisation among evaluators, evaluation products and donors. Another area where the standards have been seen as useful is in the preparation of in-house guides.

As the standards are still in a test phase, it is clear that the full extent of challenges and difficulties are not yet known. This is further indicated by the fact that six of the respondents are not sure whether they have experienced specific difficulties or challenges in regards to the standards. On the other hand, almost half of the respondents state that they have experienced some difficulties and/or challenges, while six state they have experienced none. One of the main challenges mentioned is that the standards are for general use and not for specific application. It was also mentioned that applying the standards add time to the already time constrained evaluation process.

Ten respondents included comments on specific items of the standards. Annex 2 contains further information of what the respondents who did comment thought about the different items in the standards.

### **Next steps**

Considering that 14 of the respondents to the survey are willing to share feedback on their experiences using the standards in a forthcoming workshop, it could be valuable to conduct a workshop once more recorded experiences would be available. Taking into account that most will start recording experiences in 2008, a workshop in early 2009 should be envisioned.

The objective of the workshop would be to share experiences in using the standards and agree on changes needed to reach final agreement on the standards in 2009.

At the sixth network meeting several members highlighted challenges posed by the decentralisation of evaluation functions. It was suggested that it might be beneficial to develop a 'how to' guide, based on insights from the test phase, on how to apply the standards. The workshop could address this issue, as well as other challenges and constraints, and seek ways to overcome these.

It might be useful to circulate the last part of the questionnaire, which looks at each individual standard, and the question on challenges and difficulties in advance of the workshop to obtain updated knowledge on specific items of the standards.

## Annex 1

### Answers to survey

25 respondents.

Out of which 4 are not members of the evaluation network (CTBTO (Preparatory Commission for the Comprehensive Nuclear-Test-Ban Treaty Organisation), UNICEF, WFP and the Ministry of Science and Technology in China).

### Q2. Are you currently using the Draft DAC Evaluation Quality Standards?

#### Answered 25/25 (6 commented)

80.0 % of respondents using the EQS.

All those who answered no, apart from USAID, stated that the standards they have adopted, either their own or somebody else's, are very similar to the DAC EQS, or draw upon the DAC standards.

- USAID: have adopted the American Evaluation Association's standards.
- AfDB: Some of the EQS have been adopted in the evaluation manuals and guidelines and recent evaluation policy. AfDB intends to adapt the EQS to its specific context and use in an experimental basis before adoption.
- ADB's: standards draw upon the DAC standards, use them indirectly.
- UNICEF: uses the UNEG norms and standards which are similar to DAC's.
- Japan: uses their own standards but say they cover major components of the EQS and they occasionally refer to the EQS.

### Q3. Who is using the DAC Evaluation Quality Standards?

#### Answered 18/25 (14 commented)

Central evaluation office/department	90%
Country program departments	28%
Field offices/embassies	28%
Other government departments	6%
Partner countries	0%
Consultants	56%
Civil society	0%
Other	22%

Most agencies evaluation departments who are using the EQS are also promoting the use of them to other departments and partners. Norad notes that even though the EQS have been shared with country and program departments at trainings, it does not necessarily mean that they are being applied.

### Q4. How are the EQS being applied?

#### Answered 21/25 (21 commented)

20 out of 21 who answered the question say the EQS are referred to in various stages of the evaluation process. Most (12) specifically referred to the EQS in regards to drafting ToRs, but it seems the EQS are often referred to or reflected upon in all parts of the evaluation process. Some agencies seem to refer to the EQS more than others and include them in all areas of work, while others say the EQS are taken or will be taken into account when drafting their own standards.

### Q5. Are you now or do you plan to record experiences with the EQS?

#### Answered 18/25 (18 commented)

Seven explicitly stated that they will start recording experiences in early 2008, when the EQS have been used for longer. Italy is the only one already recording experiences.

**Q6. In general, have the EQS contributed to improved evaluation work?**

**Answered 21/25 (19 commented)**

57% - yes

43% - unsure

Most of the respondents state that it is too early to say whether the EQS have actually improved evaluations, although many expect the EQS to do so. It also seems that the EQS are used rather frequently by consultants to the agencies.

**Q7. Have there been challenges or difficulties with using the standards?**

**Answered 22/25 (16 commented)**

46% - yes

27% - no

27% - not sure

A number think that the standards are not specific and clear enough. Other areas of concern are that the distinction between purpose and objective is not clear and that it is difficult to incorporate stakeholders. It is also mentioned that as the standards have not been applied fully yet challenges and difficulties are not yet fully known.

**Q8. Based in your experience, do you feel the EQS are a useful tool?**

**Answered 21/25 (11 commented)**

86% - yes

14%- not sure

The comments were very positive and the EQS are seen as helpful in setting standardised internationally agreed standards. That they are international also seems to make it easier to make consultants abide by them. The EQS have also been very useful in preparing in-house guides and approaches.

**Q9. Would you be willing to share further feedback on your experience at a forthcoming workshop?**

**Answered 22/25 (11 commented)**

63% - yes

13% - no

23% - not sure

Some suggested first half of 2008, but others thought it would be better to have it a bit later.

**Q10. Specific feedback on the quality and usefulness of the context.**

**Answered 10/25 (4 commented more)**

The following scored 50% or less on useful and complete:

- 2.2 Intervention logic and findings
- 3.3 The socio-political context
- 4.2 Assessment of results
- 4.5 Evaluation team

For a more detailed outline, please see annex 2.

## **Annex 2**

### **Q10. Specific feedback on the quality and usefulness of the context.**

**Answered 10/25 (4 commented more)**

Comments on specific content sections.

#### **1.1 The rationale of the evaluation**

There is coherence among respondents about this section with everyone agreeing that this is important information which is both useful and complete. This section does not need modification.

#### **1.2 The purpose of the evaluation**

The views are mixed on this section, some seeing it as incomplete, some as not useful and others believing it to be vital. This section might therefore require some clarification, especially since some evaluators have difficulties in distinguishing between purpose and objective (the section discussed below).

#### **1.3 The objective of the evaluation**

There is coherence among respondents about this section with everyone agreeing that this is important information which is both useful and complete. This section might still need some modification as some have difficulties in distinguishing between purpose and objective as mentioned above.

#### **2.1 Scope of the evaluation**

There might be reason to revisit this section as some believe it to be incomplete and unclear. Many, on the other hand, see it as useful and complete.

#### **2.2 Intervention logic and findings**

Although half of the respondents thought this section to be useful and complete and others saw it as vital, there might still be reason to revisit it as some thought it incomplete and not useful.

#### **2.3 Evaluation criteria**

The majority see this section as either vital or useful and complete. One agency, however, sees this section as not useful.

#### **2.4 Evaluation questions**

Although most see this section as vital and complete, there might be reason to revisit this section as a few agencies see this section as incomplete.

#### **3.1 The development and policy context**

This section is believed to be useful and complete by the majority of the respondents, only one agency thought the section to be incomplete.

#### **3.2 The institutional context**

Most respondents thought this section to be either vital or useful and complete. But two agencies thought it to be incomplete.

#### **3.3 The socio-political context**

Although most see this section as either vital or useful and complete, there might be reason to revisit it as two agencies see it as incomplete and another one as not useful at all.

### **3.4 Implementation arrangements**

The great majority of the respondents believe this section to be vital and complete, one agency, however, believe the section is not useful.

### **4.1 Explanation of the methodology used**

The majority (almost 80 per cent) of the respondents believe this section to be either vital or useful and complete, one agency however believe the section to be incomplete and another it to be not useful.

### **4.2 Assessment of results**

The majority (70 per cent) of the respondents believe this section to be either vital or useful and complete, two agencies however believe the section to be incomplete and another it to be not useful.

### **4.3 Relevant stakeholders consulted**

All respondents believe this section to be either vital or useful and complete. There is hence no need to modify this section.

### **4.4 Sampling**

The majority (90 per cent) of the respondents believe this section to be either vital or useful and complete, only one agency believed it to be not useful.

### **4.5 Evaluation team**

Almost 70 per cent believe this section to be either vital or useful and complete, while just over 30 per cent believe it to be incomplete.

### **5.1 Transparency of information sources**

The majority (almost 90 per cent) of the respondents believe this section to be either vital or useful and complete, only one agency believed it to be incomplete.

### **5.2 Reliability and accuracy of information sources**

Fewer agencies made a comment on this section. The majority, however, believed it to be either vital or useful and complete, with only one agency believing it to be incomplete.

### **6.1 Independence of evaluators vis-a-vis stakeholders**

The majority (almost 90 per cent) of the respondents believe this section to be either vital or useful and complete, only one agency believed it to be incomplete.

### **6.2 Free and open evaluation process**

All respondents believe this section to be either vital or useful and complete. There is hence no need to modify this section.

### **7.1 Evaluation conducted in a professional and ethical manner**

The majority (almost 80 per cent) of the respondents believe this section to be either vital or useful and complete, two agencies believed it to be incomplete.

### **7.2 Acknowledgment of disagreements within the evaluation team**

Near 70 per cent believe this section to be either vital or useful and complete. However, three agencies believed it to be incomplete.

### **8.1 Incorporation of stakeholder comments**

A clear majority (90 per cent) believe this to be either vital or useful and complete, with only one agency believing it to be incomplete.

## **8.2 Quality control**

Almost 70 per cent believe this section to be either vital or useful and complete. However, three agencies believed it to be incomplete.

## **9.1 Formulation of evaluation findings**

The majority (almost 80 per cent) of the respondents believe this section to be either vital or useful and complete, two agencies believed it to be incomplete.

## **9.2 Evaluation implemented within the allotted time and budget**

The majority (near 90 per cent) of the respondents believe this section to be either vital or useful and complete, only one agency believed it to be not useful.

## **9.3 recommendations and lessons learned**

The majority (almost 80 per cent) of the respondents believe this section to be either vital or useful and complete, while two agencies believed it to be incomplete.

## **9.4 Use of the evaluation**

The majority of the respondents believe this section to be either vital or useful and complete, only one agency believed it to be incomplete.

## **10.1 Evaluation questions answered by conclusion**

The majority of the respondents believe this section to be either vital or useful and complete, only one agency believed it to be incomplete.

## **10.2 Clarity of analysis**

The majority of the respondents believe this section to be either vital or useful and complete, only one agency believed it to be incomplete.

## **10.3 Distinction between conclusions, recommendations and lessons learned.**

All respondents believe this section to be either vital or useful and complete. There is hence no need to modify this section.

## **10.4 Clarity and representativeness of summary**

All respondents believe this section to be either vital or useful and complete. There is hence no need to modify this section.