



## LESSONS LEARNED FROM THE PROCESS OF THE JOINT HMG EVALUATION OF THE CONFLICT PREVENTION POOLS

September 2004

By  
Mark Keen

**PARC Project No: 172**

**The Performance Assessment Resource Centre**  
*is managed by*  
**International Organisation Development Ltd**

*2 Shutlock Lane, Moseley, Birmingham B13 8NZ*  
*Tel: (+44) 121 444 7361, Fax: (+44) 121 444 8476*  
*Website: <http://www.parcinfo.org>*

## 1. Introduction

In this short paper I have aimed to draw out lessons that can be learnt from the evaluation of the joint HMG conflict prevention pools undertaken by an external consulting team. In addition to this I have also tried to reflect on this lesson learning process and some of the issues arising from completing this assignment.

The final report and the government's response to it were published in July 2004 and were the culmination of over 400 consulting days and a less quantifiable, but highly significant amount of time, energy and resources from the client group, within a 9 month period.

As lessons have emerged, so has the possibility of new avenues of inquiry to deepen an understanding of particular issues and of possible recommendations. There are obviously limitations in what this consultancy can do, given the parameters of the output – a short paper for a potentially wide audience – and given that different aspects will be of interest to different stakeholders. As such the main focus of this paper has been to try and identify issues and lessons that are transferable to future evaluations, both of the conflict pools themselves, but also to other evaluations involving multi-stakeholder, intra-organisational initiatives or projects.

This report reflects what others have said to me, my analysis of those views, and the lessons that can be learnt and recommendations for the future. The evaluation is and has been a sensitive area, both in terms of the political context and content of what the pools are trying to do; but also as evaluation itself is often sensitive and threatening given that it is often seen to be judging groups and individuals. My view is this that these anxieties have themselves been heightened in an evaluation covering the activities of different departments who themselves are not used to working together.

It has been difficult for me to think through all the implications to stakeholders of the lessons I am suggesting, though I am aware that professional reputations could be felt to

be undermined if this is not done effectively. As such I have not mentioned individuals by name, and tried to remember a number of key points in making my conclusions:

- Who the audience is, or should be, for this output
- What is expected from the terms of reference by all the interested parties, including those who do the work outlined
- This isn't a formal evaluation, but the views are drawn from an inquiry process, as such 'evidence' should not be given any quasi-scientific validity
- The approach to this work aims to be formative, (what can we learn – do in the future) though it may be difficult to avoid all summative judgements
- What is possible within the scope of an inquiry which has a short paper as its main output

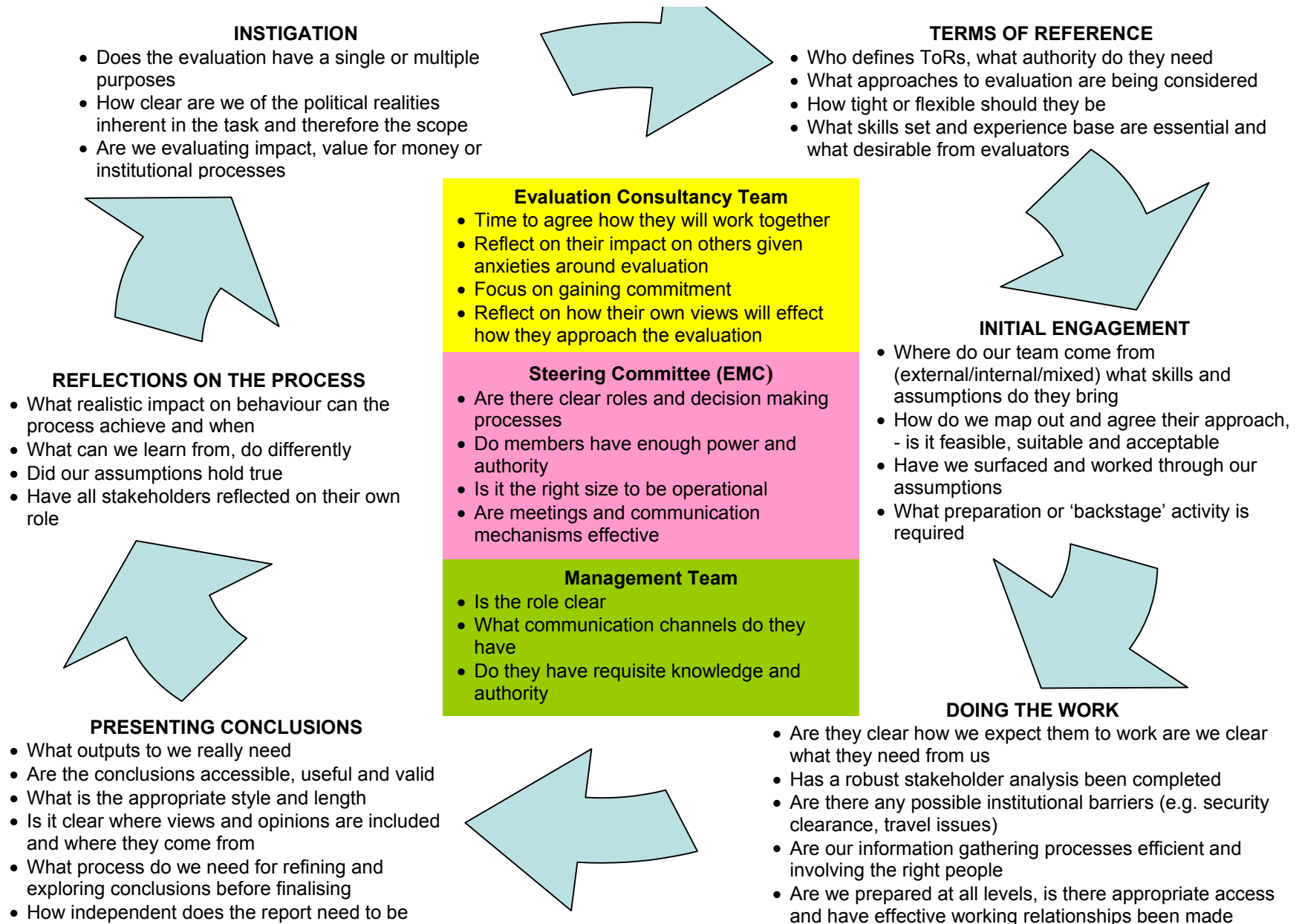
Below, in figures 1 and 2, I have tried to summarise the lessons that have been drawn out and provide a framework to guide evaluations of this sort in the future.

## 2. Approach and Methodology

The approach taken has veered slightly from the original terms of reference. In particular a questionnaire hasn't been used; instead the main inquiry vehicle has been a semi-structured interview. There were two main reasons for this:

- a) **The varying levels of engagement that different groups and individuals have had with the whole process.** This includes not only time and involvement but also knowledge and experience base. As such using a questionnaire where collation of answers is used comparatively or meaningful quantitative comparisons made would be difficult. In my judgement a richer source of data was to be gained through a semi-structured interview process bringing out what interviewees had experienced in the process and then pulling out themes

**Figure 1: A FRAMEWORK TO HELP IN THE PLANNING OF MULTI-STAKEHOLDER EVALUATIONS**



**Figure 2 – A SUMMARY OF LESSONS LEARNT FROM THE EVALUATION**

	<b>Lessons Learnt</b>
<b>Overall Impressions</b>	<p><b><i>Recognise this is a difficult task for all involved!</i></b></p> <ul style="list-style-type: none"> <li>• This evaluation was completed on time but was a difficult process partly due to it being the first evaluation done of this type in this way</li> <li>• Is it clear as to when the evaluation is primarily formative or summative – is it trying to influence behaviour or judge what has been achieved</li> <li>• A clear sound method is essential but it is important to recognise that there are alternative approaches</li> <li>• There are very different levels of experience, knowledge and interpretation of evaluation and conflict prevention both within and across government departments</li> <li>• Different expectations regarding the level of focus on evaluating conflict prevention or evaluating the pools as a government mechanism The evaluation will and has influenced policy but did not capture the ‘hearts and minds’ of all involved</li> </ul>
<b>Lessons for Evaluators</b>	<ul style="list-style-type: none"> <li>• A stakeholder analysis is a key aspect of this type of evaluation</li> <li>• Reflect on the impact your approach and style will have on those being evaluated</li> <li>• Ensure space to process how you are going to work together given complications and multifaceted nature of this type of task</li> <li>• How do you capture ‘hearts and minds’ - what is the balance needed of people being satisfied with the process yet challenged appropriately so difficult feedback can be given</li> <li>• What roles are needed within your team – do you need researchers or support staff as well as content experts and ‘evaluators’</li> <li>• Try to ensure you don’t get ‘split’ yet recognise that you work differently and will have different views</li> </ul>
<b>Formulating Terms of Reference</b>	<ul style="list-style-type: none"> <li>• Have as small a group and as simple a process as is possible</li> <li>• Recognise the cultural complexity that any team of evaluators will be facing – ensure they have time to process this</li> <li>• You can have extensive or very short terms of reference – My view is that the more complex the evaluation, the more emphasis should be placed on a negotiated process between the client group and the evaluating team. This would lead to less prominence for the terms of reference and more weight to an inception or ‘contracting’ phase, which would precede the evaluation itself</li> </ul>
<b>Agreeing who should undertake the task</b>	<ul style="list-style-type: none"> <li>• Should this be done by an external consultancy team – would a mixed internal external team more appropriate given the internal knowledge of the system being evaluated</li> <li>• There are different ‘expert’ knowledge groups who could undertake the work – chose different people and you will get a different approach</li> </ul>

	<ul style="list-style-type: none"> <li>• Try to ensure credibility and neutrality – though this will be difficult – if not possible ensure reservations and potential conflicts of expectation are surfaced</li> </ul>
<b>The 'Start Up' phase</b>	<ul style="list-style-type: none"> <li>• Good preparation is paramount – effective 'backstage' activity should enable the 'front-stage' engagement to be easier</li> <li>• A non-confrontational meeting before the work starts of all parties who will be working together is helpful, if this is not possible at least relationships to be initiated prior to formal activities so that a minimum shared understanding of what is to be done and how is gained</li> <li>• Multiparty evaluations need to try and bring together all involved parties prior to difficult issues being addressed</li> </ul>
<b>Management Arrangements</b>	<ul style="list-style-type: none"> <li>• Be careful to try and ensure the 'neutrality of the management team – if possible try and include a representative from all parts of the system though this might be difficult</li> <li>• A steering group should have authority and be as small as possible, though also inclusive</li> <li>• Meetings need to have a clear process as well as agreed outcomes</li> <li>• Recording and documenting minutes is important as is keeping a feel for informal views and opinions as they might also effect how the evaluation is received</li> </ul>
<b>'Doing the Work'</b>	<ul style="list-style-type: none"> <li>• A clear process of efficient documentation collection is important</li> <li>• Be aware of any institutional barriers that might get in the way</li> <li>• Multilayered evaluations involving processes running in parallel with some form of synthesis are difficult – case studies will be undertaken within different contexts and by different people.</li> <li>• Getting uniform outputs from multiple activities is going to be difficult – the synthesiser is likely to have a lot of pressure on them</li> </ul>
<b>Final Reporting</b>	<ul style="list-style-type: none"> <li>• Is there appropriate balance between understanding impact - 'outcomes' – and reflecting on the processes undertaken to achieve them</li> <li>• Reports are for the reader – style and presentation needs to engage as well as inform, so in general as short as is possible and with visuals as well as text if possible</li> <li>• In contrast there is also a responsibility on the reader to read – if it is a complex argument then simplifying is not appropriate</li> </ul>

and trends from these, based on an overall but loose framework.

- b) **Wariness from initial interviewees to the suggestion of a questionnaire.** It was felt that these either tend to be filled in very quickly without as much reflection and probing as in an interview, or could be 'left in someone's in-tray' until after the summer. There was also a view that those who were most animated about the evaluation process would also be more likely to fill them in thoroughly so distorting the overall picture.

In total 25 individuals were consulted, 22 in face-to-face interviews, 2 by telephone and one through an email discussion. The interviewees represented a reasonable cross-section of those involved and included at least 4 staff from each of the 3 'operational' pool departments and 2 from the Cabinet Office and Treasury. 4 of the consulting team were interviewed and 3 members of DFID's Evaluation Department.

The interviews were structured around gaining initially the interviewees impression of the evaluation as a whole and then looking separately at the stages of the process as it unfolded. The interviewees were also asked to focus on the roles taken by The Management Committee (the EMC), the Consultancy team (the Evaluators) and DFID's Evaluation Department (the Evaluation managers).

Reflecting on this approach I think it may have been possible to have combined the interviews with a questionnaire. The advantage of the interviews is that they allowed conversations to be driven not just by the loose structure applied but also by the experience and angle of the interviewee. This provided some 'rich pictures' of how different groups and individuals had experienced the evaluation but did mean sometimes conflicting perceptions. I do feel the approach gave depth but perhaps lacked the breadth that a questionnaire may have provided – in particular access to those working overseas – and also the ability for more comparative analysis. Though representatives from two country desks were interviewed - one from each pool – the

biggest gap is the lack of views from those who engaged with the evaluation in the case study countries.

Undertaking this work and reflecting on an evaluation done by others, I think it is important to recognise that methodology is important yet different approaches can be justified and challenged. The selection of a final method is both a subjective and objective process which will draw from an appreciation of the context, the skills and knowledge base of those undertaking it, the accessibility of the subject, and the resource envelope available.

### 3. Key Questions

Following the initial interviews and discussions, a number of questions have, in my thoughts, emerged which are perhaps important to address when considering cross Whitehall, multi-stakeholder evaluations of this nature:

- **Is it clear who has, driven or initiated the evaluation and why?**
- **How do different people understand the nature of evaluation, why it is done, how do you do it, and what it should aim to achieve?**
- **Who should undertake an evaluation of this nature – where should they come from, how should they be contracted?**
- **What different approaches to evaluation are there – how possible is it for alternatives to be suggested?**
- **The identification of the primary task or key evaluation question(s) – what process is needed for this to be done effectively?**
- **The process by which an evaluation team agrees its approach, prepares for the task, agrees how they will work together and what they understand their role to be?**
- **Evaluating within a large institutional system– how do you evaluate across different 'organisational cultures' as well as in different places?**

- **The process by which such a complex task is undertaken. How does it need to be managed and by whom?**

I think a number of these questions are addressed in this paper, though it is important to think more about how 'deep' an analysis is needed and how firm recommendations about what could be done differently or the same should be. My view is to try and focus on what is useful and what can be worked with, rather than perhaps open up a discussion which requires greater space than a consultancy of this nature. It may well also be that some of these questions are only important to some people, so being clear about who feels they will benefit from what would also be helpful.

#### **4. Reflection on the Evaluation Process**

One distinct aspect of this evaluation, made more pronounced by the scale and the numbers involved, is the lack of continuity of engagement throughout the process. The nature of working in Whitehall means that people often move roles frequently. Reflecting on this evaluation process few people are still in the same organisational role from when the scoping study was initiated through to the publication of the government's response to the synthesis report. There are a number of possible implications of this:

- Responses and opinions gained relate very much too each respondents own engagement or access to parts of the process.
- The views of some key individuals who played important roles are harder to access.
- The process of reviewing the evaluation becomes one of putting together a jigsaw from the various pieces of individuals' partial perceptions and experiences
- A history or 'culture' regarding certain aspects of the evaluation is created and maybe picked up by those who enter the process at a later stage. It may be possible to try and minimise this with careful documentation,

however, this relies on all those involved reading it and having a shared interpretation of both the formal and informal 'discourse'. In my view this is unlikely and therefore any evaluation of this depth should recognise and try and manage both the written account of events and the social impressions, narratives and storylines created. This 'management of meaning' is of course no easy task!

#### **4.1. Overall Impressions**

The initial question asked to those interviewed focussed on their overall experience of the evaluation. Bearing in mind the different levels of engagement there were common responses:

- It felt that it took a long time and was a laborious process. It is important to note that this doesn't refer to the professionalism of the consultants but to the process both as it was originally laid out and the reality of how it unfolded given changes and unforeseen constraints and changes.
- The terms of reference were too long, complex and open to interpretation - this led to a difficult beginning
- The drafting process of reports felt inefficient and an alternative approach would be welcomed but not sure what that would entail
- It was difficult and painful and there was conflict and disagreement

My analysis of the whole process is that the consulting team were asked to do a very difficult job, underpinned by different people within the whole system having very different assumptions about what they were there to do, and complex and ambiguous terms of reference.

In terms of time, nine months may seem a long time but it is quite short for an evaluation of this depth and complexity. For me this seems to reflect that: a lot of those involved had not experienced this type of work before; and also questions whether

undertaking this type of task, in this way is appropriate if it is felt to take too long.

The approach of the consultants was underpinned by a particular interpretation of their role based on their technical knowledge and expertise. This interpretation and the way it was communicated challenged some of the client system. A coherent, cohesive response to this interpretation did not surface in a manner (time or place) where changes in the approach put forward could be suggested and constructively worked through. This was partly due to very different levels of knowledge, experience and interpretation of:

- The processes and methodologies of evaluation
- What the possible targets of the evaluation were and how these were prioritised
- The concept of 'Conflict prevention'

Some interviewees have commented that this response needed to take place following the inception report. My view, I think, is that given the different levels of experience of evaluation amongst those involved and concerns over time and logistics this was always going to be difficult, given that a momentum had already been created. It has been clear that from that moment some groups and individuals were unhappy with the direction and approach in which the evaluation was taken. There were open disagreements and challenges and a feeling that large parts of the evaluation process were confrontational.

As such, though, the evaluation - given the vagaries, iterations and logistics of such a complex process - was completed on time and did meet the requirements of the terms of reference as interpreted. This is a significant achievement for both the evaluators and those involved in the management of the process. What perhaps it didn't do was have the 'hearts and minds' and commitment of all the stakeholders involved in the process. It could be argued that having the hearts and minds is not an imperative - e.g. people often engage more in a process when it is challenging and critical - however a number of interviewees did feel that credibility had been lost and

therefore also some willingness to 'hear' or accept some of the recommendations made.

It may be that the 'pain' some associated with the process is actually a necessary part of a difficult review of what is being done and part of the fallout of an insightful critique. It could also be that following a period of 'cooling' time from the discomfort and anger then a greater receptiveness will emerge. There is though a question as to whether that discomfort creates a resistance to the implications of those messages, - e.g. due to how they reflect on the professional reputations of certain groups and individuals - and so therefore to suggestions as to how things are done in the future. One key signifier of this was the lack of a strong press release or launch of the final synthesis report. This may affect how many people engage with the findings and any debate or possible changes that could have taken place

It is important to recognise the completion of the task within the time and resource constraints given. It is also important to note that even those interviewees who were not that positive overall, acknowledge that at least some of the recommendations have been useful and will lead to changes in behaviour (in either policy or institutional arrangements).

All interviewees agreed there are certainly lessons to be learnt in terms of how you go about setting up, commissioning and managing a joint Whitehall evaluation. There are also lessons about their approaches to evaluation in general.

#### ***4.2. Lessons For Evaluators***

This section focuses on those who undertake the evaluation. In this context it is the external consultancy team, though in future this could be an internal team or a mixed one.

In terms of best practice for the evaluators, I think there are things that it would have been desirable for them to have done differently. Some of these were within their control others restricted by time, resources,



access and the terms of reference. The major points I would pick up on are:

- **The need to do a thorough stakeholder analysis** – focusing on what different actors hopes fears and expectations of the evaluation process were. A stakeholder analysis was included in the terms of reference and would have been a significant addition to the Inception report. The fact that it wasn't done and that the evaluation process unfolded without it suggests it wasn't a high a priority for all.
- From this it might have been possible to **reflect not just on the views and agendas of key players but also on the impact the approach and style the consulting team take and its likely impact on the stakeholder group.** A good example of this concerns report writing. Some departments were critical of the style of writing and the nature and use of 'evidence'. Surfacing and addressing some of these concerns earlier might have assisted in gaining buy in to the process.
- **The need for the consultant (evaluation) team to meet as a whole team both prior to and during the evaluation itself, to work through in detail how they were going to work together. The Terms of Reference also need to build in reflective process time to do this.** The team need time to discuss key terms and methodological issues so that there is a common frame of reference. This is particularly important given the number of people in the team, their different, personal engagement styles, experience and knowledge areas and the tight timetable to complete the whole process. In interviews it was commented that the team often appeared 'split', examples given were on certain case study visits and during EMC feedback meetings. It is also important to note that given the composition of the team, the scope of the ToRs as they were and the timing that logistically it would have been challenging for them to meet up.

### **4.3. Formulating Terms of Reference**

There was significant reflection about this from members of the EMC as certainly this part of the process is one from which most feel lessons could be learnt and alternative approaches suggested.

**Who began the process of the evaluation itself and why?** - In the conversations I had few had a definitive view on this, despite the original terms of reference documentation, though those who did referred to the role of the treasury as instigators, with the purpose of the evaluation being assessing value for money. In general most were rather fuzzy about this including the treasury themselves. This lack of clarity seemed to rather muddle the process in particular as it was nearing conclusion.

**Who should be involved?** - A key element of this issue was who should be involved in this process. It was felt by many that having another external body (Kings) in this process didn't particularly add enough value given the cost incurred. The evaluators also found working from another 'expert' group's analysis difficult. If this doesn't happen, though, either a nominated internal group (or individual) has to undertake some kind of scoping study or a different approach needs to be taken. One of the lessons mentioned by an interviewee was that the Kings Study itself was done by consultants who new they weren't going to do the work. Their commitment then to it being the starting point of an evaluation rather than an interesting framework for analysis could be questioned. My view would be that a different approach is taken, where the process of engagement and what is expected is much more of an iterative process, with potential consultants taking a much less exhaustive set of ToRs which include a design phase within the evaluation itself.

**How and who determines the task?** - It is important to recognise that individual agencies and departments often have a very distinctive view and approach to the engagement of consultants and that this isn't the same for all government groups. My

experience of DFID, for example, is that there is often an assumption that the client group should know exactly what they want in some detail before they engage consultants. There are different models of engaging consultants which might be appropriate particularly for a piece of work like this that involves a complex system, multiple stakeholders, different political agendas and a high degree of ambiguity.<sup>1</sup> Other models often place a greater emphasis on encouraging the consultant and client to negotiate their terms of reference. This negotiation can be based on the consultants' expert knowledge, of how to undertake a task of this nature and what needs to be in place for it to be successful, and the clients understanding of what they want as an outcome and the resources and commitment that are available.

**Have assumptions about how you do evaluation, and what the outcome should be, been fully processed and conclusions reached?** - Agencies often take a particular approach to evaluation based on a view of 'best practice' that comes from a certain paradigmatic tradition and sectoral (e.g. for DFID, Aid and International Development) context. This has a big impact on how terms of reference are formulated. There are alternative approaches to evaluation that can be used drawing from both the expert field of 'evaluation' and also other vocational and conceptual frameworks. Other government departments don't have as clear a view or definition as DFID as to the role of evaluation and they also use different approaches (and use different language) for it. Using the DAC criteria in the ToRs seemed to push the evaluation (along of course with EvD's managing role) towards what was perceived as 'a DFID approach'. That's not to say this is a bad thing or that the DAC framework is not a useful one, but it does rather narrow down both who could do this type of work and also the approach they could take. It also puts a large degree of responsibility on DFID's shoulders as

custodians of 'evaluation' knowledge and experience. Given that this is a Cross-Whitehall initiative where one of the focal outcomes is feedback as to how well departments work together; this is probably not the best starting point.

**How flexible or exhaustive should Terms of Reference Be?** - In terms of what the ToRs should include it is important that the framework used here should not be seen as fixed and could be adapted based on the needs of different types of evaluation work. For example an approach could use similar headings but not use them all or in the same way e.g. include objectives, background, outputs, scope – and leave methods (maybe including timeframe), deliverables and conduct to a negotiated process based on a consultant proposal. In the process undertaken for this evaluation, less emphasis on ToRs and more on the generation of an agreed 'contract' at the conclusion of the inception phase. Alternatively a very different framework could be used (and is by other government departments) or a new method of engagement generated for this type of complex system evaluation.

**As few people as possible need to be involved in finalising Terms of Reference** - All respondents hoped that it would be possible in the future for less people to be involved in the process and that this would help the ToRs to be shorter and less ambiguous. For this to work though that group would need to be empowered and trusted.

**Managing different opinions and views as to what should be included is difficult in this type of evaluation. How it is done is going to have consequences** - The ToRs themselves appear very much to be a compromise and mix of a number of different views and ideas. In particular the objectives and purpose can be taken in different ways. This reflects that different people wanted slightly different things and had different perceptions as to what was possible and what the process would feel like. They also appear as 'all encompassing' and require both a high degree

---

<sup>1</sup> Examples of different approaches can be found in the consulting literature: such as the texts 'Flawless Consulting' by Peter Block, or 'Process Consultation' by Edgar Schein, ; or articles such as 'Difficult Beginnings' by Jean Neuman – full references can be supplied if required

of interpretation as well as the possibility of very different interpretations.

**Recognise the contextual complexity** - A greater recognition of the political realities, contradictions and different agendas that exist within this process. This may well be a difficult element to operationalise in formal ToRs but given that for departments to be working together effectively is a significant outcome, outlining some of the history and referring openly to some of the 'cultural differences' between how they work may have been helpful.

#### ***4.4. Agreeing who should undertake the work***

In my view it is important to reflect on what expertise is required for an evaluation of this type – Content knowledge, evaluation knowledge, context knowledge, institutional improvement knowledge? Any tendering process is likely to be based on a framework and assumptions that in themselves are likely to define what the work is and who can do it. It is important to recognise that how you define work of this nature influences who applies to do it and this is likely to be driven by the experience and knowledge base of the definers. Given that some common processes, such as the EU approach, (despite their espoused aims!) are also perceived as restrictive due to the level of resources that are needed just to complete the tender process itself, opening up work to the widest possible audience is always likely to be problematic.

**Where should your evaluators come from and what should be their knowledge base?** - The use of academics from a particular discipline area had a big impact on this piece of work. A group from a completely different knowledge base and professional pool could have interpreted the ToRs differently and still done the job effectively.

**How well do they need to know the context?** - The use of a consultant group quite closely entwined with the subject of the evaluation has some significant downsides as well as positive aspects. I accept it is often helpful to have a team who are already

familiar with the context, however one danger is that they may have a reluctance to challenge as this could affect relationships which are central to other pieces of work. They may also share some of the same assumptions as the client group about how they should work and what is possible in the future. Some conflicts of interest may be difficult to avoid - and therefore need to be surfaced and managed – if context/content knowledge is seen as essential to undertaking the work. This is likely to be exacerbated if there are a limited number of 'experts' in the area to be evaluated.

**How broad a team do you need for a task of this nature?** - In a multi-stakeholder complex system, decision-making is liable to be the product of a compromise. One of the outcomes of this was the outline for the team to include such a broad range of knowledge and expertise. This restricted who could tender and contributed to the need for such a large evaluating team and one also with little experience of working together. If such a large team is necessary then the ToRs need to build in time for that team to get used to each other and be clear as to how all their expertise should be utilised.

**How independent should the team be?** - Evaluations cause anxiety especially if they are viewed primarily as summative and so 'judgements' of performance. Given the different agendas and concerns in a multi-stakeholder setting it is important that the evaluating team be seen as credible and independent from any one interest group and as neutral as possible. If possible funding mechanisms need to reflect this and the process of redrafting reports carefully managed.

#### ***4.5. The 'Start up Phase' – confirming interpretations, agreeing scope and method***

This seems to have been the most difficult aspect of this process and the one where a lack of experience of this type of evaluation had the greatest impact. What has come out which is specific to this type of work is:

**The role of the inception phase** - one of the pivotal aspects of the evaluation seems to have been the inception phase and in particular the inception report as a mechanism for the consultants to explain how they were going to approach the task. There seems to be different interpretations as to how effective this was, though general agreement that an opportunity to really talk through different assumptions and views either did not present itself or wasn't really taken.

I think there is more work to be done in formulating different approaches to 'starting' evaluations of this type in particular the relationship between ToRs and how tightly they define the task and any inception phase where approach and interpretation is agreed. As a starting point my view would be to have either a minimal initial brief (ToR's) with a more extensive inception process; or more definitive ToRs combined with a short agreement as to how the process should unravel. My own personal opinion would be to go for the former with the inception phase based fundamentally on a thorough stakeholder analysis. There may well also be cost implications which need to be taken into account in finalising this balance.

**Good preparation is paramount** - The need for really good preparation for a process like this – in terms of 'backstage' process preparation and building up commitment to allow the 'front stage' evaluation – is key to it being supported by all stakeholders. This may also then lead to it appearing to be shorter in length. There needs to be a recognition of this in the terms of reference and to assess and surface possible differences in expectation, experience and understanding

**Multi-party evaluations need to try and bring together all involved parties prior to difficult issues being addressed** - There needs to be time for all parties to meet in a non-confrontational atmosphere. The fact that some individuals experienced confrontation at such an early stage is important and affects how the evaluation unfolds. This type of atmosphere was probably partly due to anxieties over the process itself and may not have been

helped by the style or initial take of the inception report.

#### **4.6. Management Arrangements**

Having a clear picture of what the 'correct' management arrangements for an exercise like this is difficult. Some initial reflections though are:

**Who Should Manage the Evaluation Day to Day?** - There are advantages and difficulties in having a group from within one of the main departments responsible for 'managing'. It may be the only practical mechanism but you do need to recognise what this could mean. An alternative could be a mixed team but this might be resource intensive to set up, particularly given differences in knowledge and experience of 'evaluation'.

**What is the best role of an 'EMC' or Management Steering Committee** - There is a need to reflect on the implications of being quite so inclusive, and the group becoming quite so large. Maybe the aim should be for a smaller group who have been given authority to make decisions. It might mean the need for it to be composed of quite senior staff, or to have been through a very transparent process where authority has been delegated. I think a clear picture of who is in charge is needed – so for example the role of the cabinet office and treasury to be defined, including again the requisite authority if required. I'm not sure if formal Terms of Reference are useful here or whether a more iterative process of role negotiation is more appropriate where authority is 'given' by the participating departments.

**What is the Purpose of EMC meetings** - There is a need to be really clear as to the role of EMC meetings, who should be there and what are the outcomes expected or decisions that should be made. Some views on what happened suggest that the format where the consultants presented to a large audience contributed to a confrontational process including an 'interrogation' of views and findings.

#### ***4.7. The Process of engagement – doing the work***

It is again important to reiterate that completing this work given its complexity and the differences in perception that have been expressed is a not insignificant feat. I think there are questions about the method that was used for the evaluation in particular the lack of a stakeholder analysis but also perhaps the level of activity and timing of 'data collection' in Whitehall at a Macro level. Focusing though on the approach that was used there are lessons in particular areas.

##### **A clear process of efficient documentation collection is important -**

The collation of required documents for the consultants seems to have been experienced differently by different stakeholders. The main questions seem to revolve around what is needed, why, when and from whom. Despite issues of access it does seem that a large amount was made available in a timely and useful manner. One of the administrators who helped collect documents felt that it would be helpful to have someone from the consulting team working closely with them. The consulting team also reflected on the makeup of their team and felt that maybe including one or more 'researchers' at an early stage to help in this process would have been useful.

##### **Are there any important institutional barriers that might get in the way? -**

Another key issue revolved around security clearance in particular concerning access to look for what was needed but also in terms of what 'evidence' could be quoted. This is a difficult and sensitive issue and needs clarifying.

##### **Multilayered evaluations involving processes running in parallel and then some form of synthesis are difficult –**

Its important to recognise that these difficulties will not just be concerning each separate strand but also in the linking processes and handovers that tie them together. The case studies in terms of both process and the outcomes produced; seem to have been experienced very differently. In the discussions I have had the reactions from

strategy managers have ranged from the effusive to the damning. Key points raised were:

- Preparation – more work was needed at the Whitehall level before visits were made – the team needed to get more buy in from stakeholders and be seen as credible and informed.
- There is a real need to be clear as to people's expectations. Some key individuals on the ground may not be as aware of the whole process and be defensive if they don't see the overall logic of the exercise or been a part of its development
- Case study teams need to have a balance of Country knowledge/evaluation knowledge and conflict knowledge. They need to be careful not to be seen as taking their own views, prejudgements into an inquiry process.
- Country desks/offices need to prepare thoroughly and need to be aware as to what preparation is needed
- Recognise the difficulty of getting uniform outputs when multiple activities are completed in parallel. The consulting team tried to manage this by having one main editor and an agreed case study template. This puts an enormous amount of responsibility on that individual and also may make it harder for the contributing writers who may find an imposed format difficult to work with if it is different to their own normal approach.
- They are experienced very differently – there are lots of factors to take into account and to expect: e.g. personalities, integration, fears, hopes and anxieties depending on amount of time and engagement so far
- Accessibility is an issue – if it is difficult to go out and see the required people then maybe it is important to reflect not just on how to make the best of a not ideal situation, but also whether the

chosen methodology is the most appropriate one.

#### ***4.8. Completing the Work – final reporting***

It is important to see the final report as a significant achievement given the time required and the complexity of compiling a document of this nature. The most consistent feedback, though, regarding the Synthesis report concerned the style of writing and the process by which it was finalised.

**Whose report is it and by what process should it be finalised?** – How much should the output be that of the consultants as independent evaluators and how much influenced and corrected by the client group? Focusing on the drafting process the consultants felt frustrated that some of their messages were watered down in the interest of some of those who might be perceived as being criticised. The client group questioned the process by which some of the conclusions had been reached (in particular what was based on the consultants preformed opinion and what based on evidence). The outcome does seem to have been a difficult and painful drafting and redrafting process looking at both the style and content of the report.

#### **Lessons concerning presentation: What is the appropriate density, length and style?**

In terms of presentation views have been expressed that the report:

- Is too dense – too many words and not enough diagrams or illustrations of major points (e.g. how the pools work, who reports to whom)
- Was too long and therefore of restrictive accessibility – it was seen as an academic analysis rather than as a management tool which could be used.
- Although there was a summary it is important to perhaps reflect on exactly who should do this and in what format – one suggestion would be for the summary not be written by the main author(s) but to be put

together by a copywriter or someone with journalistic training.

#### **Lessons concerning content: What is the appropriate balance between impact and process, what evidence is expected and where should it come from**

There were contrasting views as to the content of the report with most seeing some valuable insights within the analysis but not being completely happy with the whole. Some points made were:

- Not enough focus on the effectiveness of the pools as an institutional mechanism – e.g. felt the report focussed on policy and agreed activities that would have taken place with or without the pools and not enough on how the pools themselves were operating
- Is imbalanced in that it draws conclusions from specific experiences which back up the views the consultants have of what is conflict prevention

My view is that the report itself was hugely affected by the breadth of the terms of reference and their lack of clarity; and the particular interpretation taken by the consultants which wasn't really bought in to by all the recipient groups. The fact that it was also a synthesis from a number of different pieces of work also did not make the coordinating and final editing any easier. I think there are presentational issues which relate to the process being undertaken by an academic institution and also the working cultures of the different departments.

There seem to be a range of views concerning the portfolio review – some found it useful and helpful, others felt it was a bit of a 'telephone directory'. I think I probably need to ask further about the purpose and expectation of this piece of work as I'm unclear as to the role or contribution it is supposed to play.

#### ***4.9. The Impact of the Evaluation***

It is quite early to know exactly what the impact of the evaluation has been on how people are working. The HMG response

outlines initial views as to what is going to be taken on board and there are other examples where it has changed and influenced policy already (in particular Afghanistan where the conflict pool strategy is based almost entirely on the case study recommendations).

It has though had an impact and most people do seem to accept some of the recommendations. As said earlier, though, this is within an atmosphere where some individuals have strong feelings of dissatisfaction and of an opportunity not being fully utilised.

## 5. Other Lessons and Conclusions

There is a lot that can be learnt from this evaluation some of it at a practical level concerning the possible tweaking of specific elements of the process and some of a more fundamental nature.

A number of these issues are specific to an evaluation of this type and some may reflect on the nature of how HMG (in particular DFID) evaluates in general. As such it is important that other lessons, for example other work already done on evaluation methodologies and best practice, is referred to or incorporated.

There are also conclusions or lessons that I feel can be drawn which focus more on the institutional context in which the evaluation took place than the evaluation itself. These are drawn from this inquiry process so I do feel are useful to include. **From an organisation development perspective three key issues come out in terms of lesson learning;**

- **The Pools context is very political with the agendas, personalities and relationships of key senior individuals being important both now and in the formation of the pools.** These may well be key to the success of any cross-government initiative and influence what is required/possible in any evaluation. I am not sure how appropriate it is to analyse these areas further in work of this nature.

- It is clear that **the departments have different working cultures and different agendas** and that one of the government's espoused aims is to improve how they work together. It has obviously been important to some that the Pools be seen to be a positive example of joint working to ensure commitment to working at closer ties and the evaluation does reflect this. I do think **a more thorough analysis or review of what good cross-government working requires would be useful including a fuller analysis of the differences in their assumption basis and the way they work.** Again though I'm not sure this work is the right place for an analysis of that nature.
- There is **an assumption that Evaluations of this type need to be conducted by an external team.** I realise that in a lot of views of 'best practice' that is considered the optimum starting point, however given the need for credibility over such a wide stakeholder group maybe considering a mixed team of internal and external evaluators would assist in the process and help manage some of the difficult dynamics inherent in evaluation.