BENCHMARKING,
EVALUATION AND STRATEGIC MANAGEMENT
IN THE PUBLIC SECTOR

Papers Presented at the 1996 Meeting
of the Performance Management Network of the
OECD's Public Management Service

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT
Paris

50972

Document complet disponible sur OLIS dans son format d'origine
Complete document available on OLIS in its original format
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreword</td>
<td>4</td>
</tr>
<tr>
<td>Chairman’s Statement</td>
<td>5</td>
</tr>
<tr>
<td><em>Dr. Sylvie Trosa</em></td>
<td></td>
</tr>
<tr>
<td><strong>Benchmarking as a Tool for Performance Management</strong></td>
<td></td>
</tr>
<tr>
<td>Performance Benchmarking in the Public Sector: the UK Experience</td>
<td>11</td>
</tr>
<tr>
<td><em>Jeremy Cowper and Dr. Martin Samuels</em></td>
<td></td>
</tr>
<tr>
<td>The Implementation of Benchmarking from a Swedish Perspective</td>
<td>33</td>
</tr>
<tr>
<td><em>Lars Dahlberg and Christian Isaksson</em></td>
<td></td>
</tr>
<tr>
<td>Commentary: <em>Professor Helmut Klages</em></td>
<td>49</td>
</tr>
<tr>
<td><strong>Evaluating Public Management Reform -- an Illustration of Performance Evaluation</strong></td>
<td></td>
</tr>
<tr>
<td>Evaluation of Public Management Reforms in Finland: From Ad Hoc Studies to a Programmatic Approach</td>
<td>55</td>
</tr>
<tr>
<td><em>Katju Holkeri and Hilkka Summa</em></td>
<td></td>
</tr>
<tr>
<td>Commentary: <em>Danièle Lamarque</em></td>
<td>75</td>
</tr>
<tr>
<td><strong>Developing Strategic Results Indicators and Linking with Departmental/Agency Planning</strong></td>
<td></td>
</tr>
<tr>
<td>Strategic Management in Government: Extending the Reform Model in New Zealand</td>
<td>81</td>
</tr>
<tr>
<td><em>Alex Matheson, Gerald Scanlan and Ross Tanner</em></td>
<td></td>
</tr>
<tr>
<td>Commentary: <em>Dr. Gudfinna Bjarnadottir</em></td>
<td>101</td>
</tr>
<tr>
<td><strong>The Contract Agency Approach as an Illustration of Performance Management</strong></td>
<td></td>
</tr>
<tr>
<td>The Contract Management Project in Denmark</td>
<td>105</td>
</tr>
<tr>
<td><em>Peter Kjaersgaard Pedersen, Hanne Dorte Sørensen and Jakob Buhl Vestergaard</em></td>
<td></td>
</tr>
<tr>
<td>Commentary: <em>Professor Anne Drumaux</em></td>
<td>115</td>
</tr>
</tbody>
</table>
FOREWORD

This publication contains the papers and commentaries presented at the 1996 activity meeting of the Performance Management Network of the OECD’s Public Management Service. Representatives from 26 of the 29 OECD Member countries, the European Commission, and other invited experts participated in the two-day meeting held at the OECD Headquarters in Paris on 4-5 November 1996. The meeting was chaired by Dr. Sylvie Trosa of the Department of Finance in Australia.

The purpose of the meeting was to discuss contemporary developments in performance management taking place in the national administrations of the OECD Member countries. Performance management, in its broadest sense, covers reforms introduced in order to bring about a greater focus on results, particularly with respect to efficiency, effectiveness and service quality. Performance management involves both the measurement of performance and the mechanisms by which performance is made to “count” for something in government. These mechanisms include, for example, performance contracting, developing linkages between performance indicators and the budget process, and other forms of performance-related agreements.

Each session included a paper and presentation on the experiences of a selected country with respect to a particular area of performance management, followed by a commentary on the paper by an expert from outside of government. These presentations then set the stage for extensive discussion and debate by the participants. The session topics were:

- **Benchmarking as a Tool for Performance Management**;
- **Evaluating Public Management Reform -- an Illustration of Performance Evaluation**;
- **Developing Strategic Results Indicators and Linking with Departmental/Agency Planning**;
- **The Contract Agency Approach as an Illustration of Performance Management**.

The meeting was organised by David Shand (Acting Head of Programme), Marjukka Ala-Harja, Sigurdur Helgason and Christine Lidbury. This publication was compiled and edited by Christine Lidbury. The views expressed in the papers and commentaries are those of the authors and do not commit or necessarily reflect those of national authorities in the countries concerned. The report is published on the responsibility of the Secretary-General of the OECD.

Additional information on the Public Management Service’s work and publications in the area of performance management can be found on our Website at http://www.oecd.org/puma/work.htm#mgmtres.
CHAIRMAN’S SUMMARY

by Dr. Sylvie Trosa
Department of Finance, Australia


The conference opened with excellent presentations on the UK and Swedish experiences with benchmarking. They showed that process benchmarking and quality initiatives tend to be more widespread and better linked within agencies oriented toward continuous improvement and re-engineering. Moreover, they showed that where benchmarking is not properly linked to continuous improvement approaches, there is a danger that staff will not take ownership of changes. Thus, benchmarking should be seen as part of a new performance culture, and not as a one-off event.

Most countries benchmark against the private sector because it is seen to offer the most challenging standards, as well as greater credibility with the public. No countries benchmarked solely against comparable organisations in the public sector. The trend in benchmarking is to find and take up the most challenging benchmark standards.

The two main types of benchmarking -- i.e. process benchmarking and results benchmarking -- are increasingly seen as complementary methods to be used in tandem. Results benchmarking can be used to identify discrepancies in results (e.g. one prison using double the resources of another), and process benchmarking can go on to explain why the discrepancies exist. However, process benchmarking without results benchmarking can become overly inward-looking, leading to a focus on enhancing processes for their own sake without a check on whether or not the changes are particularly relevant for clients and stakeholders. There was consensus among countries that taking a comprehensive view of benchmarking resulted in a number of benefits, including raising challenging questions about performance, increasing knowledge of the whole range of stakeholders, and improving work processes.

In general, benchmarks are not used as tools for resource allocation -- at least not officially. This may be because the link between benchmarking and resource allocation is complex. For example, giving more resources to the best performing organisations may widen the gap between them and poorer performing organisations without addressing the real performance issues. On the other hand, not rewarding the best performers might be discouraging, and even serve as a disincentive for ongoing performance improvements. In some cases, it may be preferable to use resource allocation to benefit those poorer performing organisations in exchange for a commitment to improved performance targets. Thus, there is not a one-way link between the results of benchmarking and resource allocation decisions. As some participants noted, “don’t let figures speak for themselves.” Benchmarking is a good “trigger” to start the analyses of why there are gaps and discrepancies over results, but it is less effective as a diagnostic tool for prescribing how to deal with those results. Benchmarking is also a useful devise for engaging public pressure, whereby those with “bad” benchmarks feel obliged to raise their status. Benchmarking for purposes of resource allocation may be easier at a decentralised level, e.g. in regional
offices or single-purpose organisations (as in France), rather than at the macro level, e.g. between large organisations or divisions where activities are more complex and difficult to compare.

Some of the current benchmarking models have been criticised for being too inward-looking, focusing mainly on increasing organisational performance but taking the viewpoint of the client (including the Minister) only marginally into account. For example, data is often collected via instruments such as client satisfaction surveys. Yet such surveys do not reveal anything about client expectations for the service because they assess satisfaction with given services, but not client’s needs for services in general. Nor do they reveal anything about a service’s effectiveness in implementing government objectives or social or economic impact. An overly inward-looking approach to benchmarking could lead to organisations that are perfectly managed but do not achieve the right policies. To avoid such problems, for example, a project to benchmark local authorities in the UK used consultations with community organisations to set benchmarks and then linked the benchmarks with policy priorities.

Benchmarking exercises will yield the best results where managers recognise the links between processes, outputs and outcomes. This is because results benchmarking will not provide all of the answers for why an organisation has achieved (or failed to achieve) certain levels of performance. Understanding what is behind differences in performance levels is as critical a task as uncovering the differences themselves. Similarly, process benchmarking will yield the best results where managers avoid becoming lost in detail, and instead focus on key tasks for achieving outputs. Process benchmarking is improved when it is aligned with, and facilitates, programme and organisational objectives. The diagram below illustrates this process.

Benchmarking appears to require a fairly strong central impetus -- i.e. when there is no competition or other incentive, organisations are reluctant to benchmark without a central directive. Most countries believed that implementation of benchmarking should be devolved, but that the commitment to it had to come from a clear central framework focused on assessing outcomes rather than focusing on methods.

International benchmarking is considered as a priority by most countries. An international focus is necessary to capture public sector activities where the only possible benchmark comes from similar public service activities in other countries.
The conference also addressed the relationship between benchmarking and market testing. A “devil’s advocate” viewpoint might be that it is quicker to market test and see who is the best than to launch costly benchmarking initiatives. Nevertheless, increasingly there is a view that benchmarking is a better approach for activities that are under a certain financial threshold, and for activities that will clearly remain in-house. Thus, market testing should not be used as a tool to bring efficiency and market discipline into the public sector where there is no desire to outsource the activity or to ensure a level playing field between the public sector and the private sector.

The discussion on evaluation opened with a presentation on Finland’s experiences in evaluating the results of its public sector modernisation processes. The commentary following the presentation compared and contrasted the similar experiences of France in assessing its reforms. Both countries had found similar difficulties, namely:

- the absence of comparable data on the situation before modernisation, resulting in difficulties in measuring progress;

- political sensitivity about having negative or critical conclusions published.

Both countries also acknowledged similar lessons. First, an evaluation project is easier when launched at a favourable political time: for example, just after an election, when a new government might be more willing to accept criticism, as opposed to at the end of a government’s mandate, when it may be least likely to accept criticism. Second, internal evaluations (i.e. by public servants for public servants) are better accepted within the public service, while external evaluations (i.e. ‘outsiders’ questioning the appropriateness of public services or processes) are often seen as challenging or threatening, making it more difficult to build consensus among stakeholders. This raises an issue of balancing the costs and benefits -- i.e. the need for the evaluation to be accepted by the public service versus the capacity of outsiders to ask the tough questions. It also shows the primary importance of strategically managing stakeholders’ expectations (e.g. clients, the general public, public servants, politicians, evaluators), particularly as they seldom have the same interests and expectations.

The conclusions about the use of evaluations were not very optimistic. They showed that governments are still very reluctant to publish evaluations when they cannot completely control completely the results. Moreover, they showed that public servants remain reluctant to accept any form of expertise outside their own on the grounds of “efficiency” and expediency. The public service often fails to accept that an internal evaluation might simply “reinvent the wheel” of issues and problems already largely known to those involved.

**Contract Management and Strategic Planning**

*The new deal between ministers and public servants*

Presentations on contract management in Denmark and on developing strategic results indicators in New Zealand sparked extensive discussion on the issues behind contractual relationships within the public service, particularly those between ministers and public servants. In many countries it was said that ministers push hard to ensure that departments do what they expect, making the lack of trust between political leaders and the career public administration quite obvious and, due to the lack of trust, leading to more and more interventions by ministers into micro-management of the agency.
In New Zealand, they call this blurring of the traditional Westminster model a “purple zone” -- i.e. the zone where the boundaries overlap between the need for an independent and fearless public service to advise politicians, and the need for the support of public servants on political priorities. This blurring makes the management of relationships with ministers more complicated. Their conclusions addressed the need to take ministers’ concerns better into account, explain actions to politicians when possible (perhaps even provide some training if necessary) and involving them in projects at an early stage -- all while not losing the capacity to give frank and unbiased advice. It was also acknowledged that the “logic” and priorities of ministers and that of public servants are not always the same, and that public servants thus need to make efforts to understand ministers’ concerns and address them using less bureaucratic language.

From legal contracts to the ethos of contractualism

New Zealand has become increasingly conscious that output-based budgeting often can be too burdensome and detailed and, moreover, pushes agencies to focus solely on their own outputs and aims at the expense of achieving government-wide objectives. This perspective was reinforced by Allen Schick’s report evaluating New Zealand’s large-scale reforms in the public sector. In order to minimise these problems, New Zealand has built a management model based on an iterative series of priority-setting activities starting with strategic result areas (i.e. government-wide priorities), which are further broken down into key result areas (i.e. agency-specific outputs). These are used to form the basis for chief executive performance agreements (i.e. accountability for results), and thus offer strategic direction for agency activities and resource allocation.

Contracts offer a way to supplement the over-formal nature of targets and indicators. Contracts place more emphasis on “managing relationships” than on simply collecting quantitative information about achievements. For example, participants discussed at length the importance of bolstering performance contracting with strategies for conflict resolution, leadership enhancement, and other management tools that focus on the people-side of performance. There was consensus that a “learning organisation” paradigm has replaced the programme management paradigm seen in the 1970s.

Several countries emphasised that, while contractualism is gaining new attention in the field of performance management, and is finding new means of expression as in New Zealand, it is a very old practice. For example, in France there have been performance contracts between the State and government business enterprises since the 1940s.

In discussing the use of contractualism as a model of performance management, there was consensus that where the public service makes contracts with itself (as in performance contracts between central and line agencies, or ministers and civil servants) the result is generally a quasi-contract or agreement -- i.e. two parties commit themselves to a set of rights and obligations in a formalised written document, but one that is not “legally” enforceable. Within the public service there are very few cases of “real contracts.” One explanation for this was that “internal” performance contracts are used to manage the various working relationships and provide incentives for improved performance, rather to establish a more formal enforcement/sanction-based system. The opposite is true for contracts between the public and private sectors (as in contracting-out activities), which are based on formal, legally binding contracts and procedural sanctions, thus making them quite different in purpose and content from internal performance agreements.

---

The Danish presentation on contract management highlighted the idea that sanctions are not always a necessary element in internal performance contracting. The fact that such a contract is published and serves as a tool for negotiation can be sufficient incentive for both parties to fulfil their contractual obligations. Moreover, the public visibility of breaching a contractual relationship within the public service is thought to be as effective an incentive as process or legal-type sanctions.

For most countries, contractualism replaces the paradigm of devolution. The limits of the devolution paradigm are becoming apparent as it shifts from a focus on managerial flexibility to a focus on almost one-sided target-setting: that is, where agency managers have come to feel that they are fairly autonomous in promoting their agency-specific objectives, rather than broader government-wide objectives. Autonomy should always have a counterpart in ensuring that government’s key priorities are also being achieved. Contracts, as the counterpoint to devolution, highlight the need for a balance between rights and duties for both parties in the contract. Furthermore, the word “contract” implies the desire to get away from a command-and-control mode of managing in exchange for a more balanced approach based on negotiation and consultation.

The Learning Organisation

The strongest cross-cutting message from the discussions on selecting and evaluating performance management models was the need for countries to develop a change-management strategy. Traditionally, countries might have balanced between two management models -- one model promoting a single (government-wide) management solution imposed across the board to ensure consistency and expediency in implementation, and a second model of letting the managers manage within the generic framework allowing each agency to choose its own reform tools and when and how to implement them. This second model has the benefit of maximising staff ownership of changes, but can result in a very patchy situation where some agencies are much more advanced than others. Currently, the move toward contractualism offers a third and alternative model. However, the feeling around the table was that there is no single model applicable to all countries or even to all agencies within a single country, but that sufficient analysis and cross-fertilisation mechanisms nevertheless have to be put in place if agencies do not want to “reinvent the wheel” or make the same mistakes as others who have tried these methods before them. The model one chooses has to fit with:

− the aims of a reform;
− the differences between the missions and “professional identities” of the agencies involved;
− the individual organisational cultures;
− the legal framework.

Reforming the performance and accountability systems in public service organisations is being increasingly recognised as a learning process rather than a quick fix. While it is a process that evolves and changes over time, most countries still call for a guiding “framework” setting out the fundamental principles of each element of the reform. For example, should service delivery be outsourced? If so, why? Is a purchaser-provider model applicable to all situations? If not, to which situations? More discussion is needed on how to develop such a framework and create a consensus around it. The papers and commentaries that follow offer many insights and raise thought-provoking questions about these contemporary issues of performance management and reform.
PERFORMANCE BENCHMARKING IN THE PUBLIC SECTOR:
THE UNITED KINGDOM EXPERIENCE

by Jeremy Cowper and Dr. Martin Samuels
Next Steps Team, Office of Public Services
Cabinet Office, United Kingdom

Introduction

There has been significant reform of the public sector in the UK since 1979. A wide range of initiatives has affected every area of activity, including central government, the National Health Service and local government. A common feature of these initiatives is the drive to improve value for money.

Amongst the more important individual initiatives have been the creation of Next Steps Agencies to undertake many of the executive functions of central government, market testing and compulsory competitive tendering of many operational functions throughout the public sector, the transfer of previously public businesses into the private sector where this was feasible, and the creation of internal markets, such as in health and education, to replicate competitive environments. This quest for improved value for money has led to the development of a range of efficiency tools. Within this context, “benchmarking” is one of the several tools which are increasingly recognised as particularly valuable.

“Benchmarking” as an efficiency tool is based on the principle of measuring the performance of one organisation against a standard, whether absolute or relative to other organisations. It can be used to:

- assess performance objectively;
- expose areas where improvement is needed;
- identify other organisations with processes resulting in superior performance, with a view to their adoption;
- test whether improvement programmes have been successful.

Benchmarking can be effective at all levels of operation, from the conduct of individual processes, such as invoice handling, to the operational performance of organisations with tens of thousands of staff, such as a welfare benefits delivery agency.

Benchmarking as a family of techniques has three main aspects. Using the terminology adopted in a paper presented to the November 1995 PUMA Performance Management meeting by Dr Sylvic Trosa and Suzanne Williams in their paper on the Australian experience of benchmarking, these are:

- Standards benchmarking -- setting a standard of performance which an effective organisation could be expected to achieve. The publication of a challenging standard can motivate staff and demonstrate a commitment to improve the service provided. Information on an organisation's
performance against the standard can be used as a monitoring tool by its principals -- ministers or councillors.

- Results benchmarking -- comparing the performance of a number of organisations providing a similar service. In the public sector, this technique can serve to allow the public to judge whether their local provider makes effective use of its resources, compared to other similar providers. In the absence of the competitive pressures which operate in the private sector, this can provide a significant incentive to improve efficiency.

- Process benchmarking -- undertaking a detailed examination within a group of organisations of the processes which produce a particular output, with a view to understanding the reasons for variations in performance and incorporating best practice.

The emphasis of this paper is on the use of benchmarking within UK central government. However, it also examines examples of the use of all three aspects of benchmarking within the public sector in the UK, seeking to outline the philosophies behind the approaches used, to give brief details of the methodologies adopted, to highlight issues and problems which have arisen, and to consider the extent to which the results of benchmarking have been used in decision making.

**Philosophy**

At the heart of many of the public sector reform initiatives in the UK has been the philosophy that value for money can be best achieved by a separation of roles between those who set the policy and those who deliver it -- between the purchaser and the provider. This shift to a more contract-based system gives managers freedom to manage the operational delivery of services, within the policy and resources framework established by the political authorities. Greater clarity of roles and responsibilities has proved to be a key feature in facilitating improved performance.

Within this framework, however, managers must have incentives to use their freedoms to improve the efficiency and effectiveness of their organisations. In the private sector, this incentive is supplied by the competitive nature of the market -- low quality organisations will go out of business. The majority of public sector services, by contrast, do not operate in a competitive environment and therefore do not experience this pressure to improve.

A significant strand of the UK Government's drive to improve the performance of the public sector has been the creation of alternative means of replicating the pressure to improve which exists in the private sector. This has included requirements set by central government for local authorities to put certain activities out to competitive tender and for all public services to consider areas of work for transfer out to the private sector.

As experience with these initiatives grows, the focus is shifting from specifying the use of particular efficiency tools towards allowing organisations to select the techniques most appropriate to their particular circumstances, though they may be challenged to justify their choices. This freedom, however, is within the context of moving towards measuring and publishing organisations' performance, as a means to identify good practice and encourage the pursuit of improvements -- i.e. benchmarking. Through this approach, the UK seeks to achieve continuous improvement of public services, while retaining public accountability for service provision.
Standards Benchmarking (1) -- Next Steps Agencies

Background

In October 1995, the Deputy Prime Minister, Michael Heseltine, announced his intention to benchmark the performance of central government Next Steps Executive Agencies against both the private sector and, in due course, public services in other countries.

It is expected that the restructuring element in the Next Steps project, launched in 1988, will shortly be completed. Currently, 375,000 civil servants -- 71% of the total -- work in the 125 Next Steps agencies, or organisations operating fully on Next Steps lines. Benchmarking performance on such a scale is a very major task. It was therefore decided to begin the process with a pilot exercise to test one possible approach.

Methodology

Next Steps Agencies perform most of the executive functions of central government, ranging from the payment of welfare benefits to the provision of forensic services to the police, and from preserving official records to protecting fish stocks. As a general rule, each Agency is unique within its own area of business, making direct comparison between the performance of different Agencies -- results benchmarking -- problematic.

Following research and consultation with Agencies, it was decided that the pilot exercise should seek to measure the total activity of Agencies. By stepping back from the specific activities undertaken by each Agency, it would be possible to examine performance in key areas, such as people management, which are common to all organisations, regardless of their area of business. This approach has the advantage of allowing comparisons to be made between Agencies and complements work to assess better the tasking and the business results of Agencies which is the subject of a parallel Government initiative.

Tenders were invited from consultants to undertake the pilot exercise. The proposals received were broadly divided into those which intended to develop a methodology specifically for benchmarking Next Steps Agencies and those which applied an existing model. Since a key objective of the exercise was to facilitate benchmarking with the private sector, an existing methodology was selected, allowing comparison with existing data on other organisations. The tender which met the requirements most fully was the proposal from the British Quality Foundation to use its ASSESS RapidScore product to undertake standards benchmarking.

The British Quality Foundation (BQF) was established in 1992 to promote the principles of total quality management throughout British industry. One of its key roles is to act as the intellectual custodian within the UK of the Business Excellence Model (BEM), developed by the European Foundation for Quality Management (EFQM) during 1988-1991. The BQF uses the model as the basis for the UK Quality Awards, which recognise British businesses of outstanding quality. The model setting out its nine criteria is reproduced at Annex A.

A full assessment of an organisation's performance against the BEM is a significant undertaking, requiring considerable effort and resource from senior management. The BQF has recently developed its ASSESS RapidScore product as a simple approach to assessment, which is designed to encourage businesses to use the BEM and so become involved in the culture of continuous improvement which it represents.
RapidScore provides the means for organisations to use self-assessment to derive measurable and comparable data on their performance against the BEM. It contains 91 questions, linked to the nine criteria of the model, which explore the approach, deployment and results achieved in key areas. This simple methodology allows realistic assessments to be made while reducing the demands on senior management time. In order to improve calibration of scores against the model, consultancy support has been provided to the Agencies undertaking the self-assessments.

A total of 30 Agencies, covering almost 200,000 staff (Annex B), will undertake self-assessment using RapidScore. Since involvement in the pilot was voluntary, they are not a representative sample, tending to include a higher proportion of the largest Agencies. Several other Agencies with experience of using the BEM are associated with the project. The first Agencies began the self-assessment in June 1996, and the last will complete by the end of the year. The results for 26 Agencies are summarised in the chart in Annex E. The summary in Annex E gives the highest, lowest and average scores recorded in each of the nine categories and compares them to the standard set by the UK Quality Award winners. Since the project is still underway, results and interpretation are only preliminary.

In general, the scores achieved indicate that Agencies are performing to a respectable standard, with few scores giving obvious cause for concern. The best scoring Agencies are within sight of the UK Quality Award standard in due course. However, it should be borne in mind that the scores are the result of self-assessment, facilitated by external consultants, and may therefore not be as robust as scores achieved by award winners, which are externally validated.

In running the pilot exercise, care was taken to ensure that Agencies went through the same process as do private sector organisations. This allows a direct comparison to be made between the results from the public and private sectors. Although the BQF does not release the results from individual companies for reasons of commercial confidentiality, it is possible to compare the average scores under each criterion held on the BQF database for the private sector against those for the Agencies. Again, scores are compared to the standard set by the UK Quality Award winners. This data makes it possible to identify those areas where Agencies appear to be performing particularly well, as well as areas where further improvement appears necessary.

In comparison with the private sector, Agencies scored well in the areas of customer satisfaction, business results, policy and strategy and the management of service delivery generated by agency status has enabled agencies to develop more robust business strategies, with particular emphasis on the effective management of their financial resources. Agencies, encouraged by the Citizen’s Charter, have also become more responsive to their customers and are more sensitive to external factors. With this has come a greater sense of corporate identity and a culture which strives for continual improvements in efficiency and effectiveness.

Agencies scored noticeably better than the private sector average in the criterion covering customer satisfaction. This suggests that efforts over recent years to improve the effectiveness of public services and raise the importance of the public as customers for services have borne considerable fruit. However, there is some evidence to suggest that the current high levels of achievement in this area are, in part, the result of starting from a comparatively low base, and that Agencies may not be able to improve their services still further as quickly as customer expectations rise. If so, levels of customer satisfaction may over time become more in line with private sector experience.

The Agencies tended to score rather less well than the private sector in the areas of leadership, human resources, processes and employee perception. The results indicated that there was still considerable room for improvement in the areas of internal communications, including feedback from both...
managers and staff, human resource planning, and the operational management of processes and their interfaces.

This may in part be a consequence of the fact that the civil service has only recently moved away from a system under which most key decisions affecting staff management, such as pay and grading, were taken centrally, with limited input from individual business units. It would not be surprising if moves to devolve control over these areas of business down to individual Agencies appear to have resulted in a degree of uncertainty, as agencies start to include the new aspects in their planning systems. However, as the process becomes more familiar, it can be expected that the freedom to design human resource systems which meet the particular needs of the business of the individual Agency will result in significant improvements in performance. It can also be expected that the commitment for all civil service organisations to achieve “Investors in People” accreditation by the year 2000 will have an impact on results here.

Issues and problems

The reaction from the Agencies taking part in the pilot exercise has been largely favourable. In general, Agencies have found it valuable to assess their position against both the standard set by the UK Quality Award and also the results achieved by the private sector. Their experience suggests that much of the approach of the private sector applies equally to public services, though certain issues (noted below) have arisen.

The process of self-assessment has provided benefits to Agencies both in terms of improving their knowledge of themselves and of highlighting areas of difference between the disciplines of the public and private sectors. Specific benefits of benchmarking identified so far are that it:

- provided Agencies with an insight into private sector best practice, as described in the model, and caused key issues affecting modes of operation to be raised for further discussion;
- focused the management teams on their business in total, providing an opportunity to discuss and debate key issues and their linkages and interdependencies;
- provided a process for managers to map their progress to date and plot where they needed to be in future in relation to their plans;
- provided the basis for identifying issues affecting Agencies as a group and more specific local issues that could be developed via inter-Agency contact and sharing of best practice.

Certain key issues emerged with regard to the extent to which the BEM is applicable to the public sector:

- the extent to which customers can drive the policy of the Agency. The model gives this a high value, but Agencies carry out Government policy, and Next Steps arrangements must deliver the balance between the needs of taxpayers, customers and staff, inherent in the priorities set by Ministers.
- the market context in the model has also highlighted the problems experienced by Agencies in attempting to focus on this competitive issue. Direct competitors are in many cases difficult to identify and the focus for most Agencies is on achieving objectives efficiently rather than on out-performing competition.
the extent to which an agency in itself impacts on society in carrying out its function has proved complex to evaluate. For most private sector organisations, their broader reputation within society can have a significant effect upon the willingness of customers to buy their product. This area can, therefore, be of considerable importance for their business success. However, this pressure applies in a different nature to those agencies whose customers may have little choice as to whether to use the service.

whereas the purpose in benchmarking agencies is to facilitate their performance rather than draw in evaluation of policy matters, the BEM is designed more holistically. Agency targets are based on outputs rather than outcomes, in line with the philosophical split between policy and operations inherent in Next Steps.

These areas of comment indicate that the environment in which Agencies operate, carrying out policies set by Ministers, who are accountable to Parliament, is in certain important respects different from that in which private sector organisations exist.

These issues in applying the BEM to the public sector, coupled with the complexities of calibrating scores accurately against the model, are proving challenging. Greater experience with the model will also provide assurance or re-evaluation about the robustness of scores generated through self-assessment. Accurate calibration of scores is essential if the process is genuinely to benchmark Agency performance against a set standard, thereby allowing scope for comparison with other organisations, rather than merely to give an indication of areas of relative strong and poor performance within a single Agency.

Use in decision making

Following the self-assessment, Agencies identify key areas for improvement and develop appropriate action plans. Many of the improvement actions to date reflect the areas where the self-assessment scores identified weakness in performance, in particular “communications”. A number of Agencies have used the results of the benchmarking as a catalyst to drive forward initiatives to improve links both between staff and management, and between the Agency and its key customers.

An encouraging sign from the improvement plans is that, despite the generally high scores recorded in the areas of Customer Satisfaction and Business Results, Agencies recognised that scope remained for raising performance still further, while consolidating the achievements already recorded. Agencies focused on improving the link between their overall strategy and its implementation throughout the organisation by means of business planning.

Agencies have proved ready to accept the results of the benchmarking exercise, which have generally revealed new aspects of known issues, rather than uncovering previously unsuspected weaknesses. This indicates that the greater focus on improving the management and delivery of the executive functions of government -- the main goal of the Next Steps initiative -- has taken root within Agencies. The managers and staff of Agencies have demonstrated a commitment to the principles of increasing the efficiency and effectiveness of their organisations and are seeking to adopt a culture which aims at continuous improvement. Nevertheless, it is clear that there is still considerable scope for improvement in the achievement and measurement of output performance.
International comparisons

An important element of the Deputy Prime Minister's strategy relates to the potential to benchmark Next Steps Agencies against public services in other countries. One reason for selecting the BQF's tender, and using the BEM, was that the results from the initial pilot exercise should in due course be comparable with data gathered by the EFQM for the European Quality Award.

We understand that the EFQM is working with the European Commission to adapt the BEM for use in the public sector. During the course of the pilot exercise with the BQF, certain difficulties have been found in applying the existing model unadapted to the particular circumstances of the public sector. The UK therefore has considerable interest in the progress of the EFQM's work.

Standards Benchmarking (2) -- The Citizen's Charter

Background

The Citizen's Charter, launched in 1991, is a ten-year programme which aims to raise the standard of public services and to make them more responsive to the needs and wishes of users. It covers bodies in the public sector, such as central government Next Steps Agencies and units of the National Health Service, and also the private regulated monopoly utilities supplying electricity, gas and water.

The Charter is based on six key principles:

- standards;
- information and openness;
- choice and consultation;
- courtesy and helpfulness;
- putting things right; and
- value for money.

Much of the work to implement the Charter's principles has focused on the area of results benchmarking and is described in the section covering that topic. Two initiatives, however, fall under the category of standards benchmarking -- Charter Mark and the six service standards for central government.

Methodology

The Charter Mark Award Scheme, launched in January 1992, rewards excellence in delivering public services. It is open to all public service organisations that serve the public directly. Organisations receiving the award must provide evidence of their performance against the six Charter principles, show year-on-year improvements in the quality of their services and demonstrate that their customers are satisfied with the services they provide. They must also have plans to introduce at least one innovative enhancement to their services which can be put into effect without increasing the cost to taxpayers or consumers. Independent assessors visit short-listed organisations.
To be considered for an award, an organisation must apply to the Citizen's Charter Unit in Cabinet Office (Office of Public Services). Since 1995, members of the public have been able to nominate organisations for the award, and a publicity campaign has been run each year to generate awareness of the scheme. In 1995, 4,010 nominations were made; in 1996, 15,132 (covering over 5,000 different organisations).

From a total of 2,707 applications under the scheme, 451 Charter Marks have been awarded, and a further 298 organisations have been highly commended. Since the standard of service provided by an organisation may vary over time, Charter Marks are awarded for a period of three years only, and renewal is dependent on a fresh application and assessment. The importance of this measure is demonstrated by the fact that to date only 25 of the 36 organisations whose Charter Marks have expired have been successful in having the award confirmed for a further three years. It is hoped to expand the programme to at least 2,000 Charter Mark holders by the year 2000, while maintaining its current high standards.

The second area of standards benchmarking resulting from the Charter is the adoption of six service standards by the whole of central government, including both Next Steps Agencies and the centres of Departments, as from 1 April 1997. The standards cover:

- Targets for answering letters;
- Keeping to agreed appointment times;
- Providing clear information about services and a telephone contact number;
- Regular consultation with users about the services provided;
- Complaints procedures;
- Making services available to everyone.

For the first time, the performance of all government Departments and Agencies will be measured against this single set of indicators and the results published. The Citizen's Charter Unit is consulting service users about the standards and will revise them in the light of users' views.

Issues and problems

The organisations which have applied under the scheme during its first five years represent only a small proportion of the estimated 75,000 public service organisations across the UK that are eligible to apply. The challenge for the next five years will be to bring Charter Mark to bear on the vast majority of public services.

An expansion of the scheme would bring its own problems. The existing assessment process, which is geared to receiving up to 1,000 applications, lasts three to four months and requires a considerable effort and cost to run. Alternatives, such as turning Charter Mark into an accreditation scheme or introducing a rolling programme of applications and assessments, are being considered for the longer term.

The six service standards apply to all central government Departments and Agencies, but not to other services, such as schools and hospitals, where there is much less direct Ministerial control. One of the challenges for the future is to continue the process of getting these organisations to set standards for their services. There are already more than 10,000 local charters, covering services as diverse as doctors' practices and fire services. But for benchmarking to be successful, organisations need to set the same standards so that meaningful comparisons can be made.
Use in decision making

The Charter Mark Award scheme has, for the first time, established a standard of performance against which the achievement of individual public service providers can be benchmarked. The creation of this standard, which is sufficiently realistic as to be attainable yet sufficiently challenging as to require considerable effort for its attainment, has served as the impetus for the adoption of similar schemes within a number of the larger service providers, such as the Benefits Agency's Quality Award scheme.

Charter Mark enjoys high level support from Ministers and has therefore served to raise the profile of customer service within the public sector. The existence of an award to recognise achievement in this area has undoubtedly acted as an incentive for managers to examine the performance of their organisations critically with a view to making improvements. It may be expected that the new service standards will have a similar effect.

International comparisons

The nature and scope of the organisations covered by the Citizen's Charter are such that there are considerable practical difficulties in seeking to make direct comparisons with services in other countries, though this is an area of interest.

Results Benchmarking -- Local Authorities

Background

The previous sections have been largely concerned with benchmarking the activities of central government. In addition, here are initiatives to measure the performance of individual schools, hospitals and ambulance trusts, local authority and emergency services, which allow organisations to benchmark their performance against other, similar, services. The results are published in performance tables that are made available to the public. As an example of how the process works, this section looks at the programme of results benchmarking of the performance of local authorities, including police and fire services, in England and Wales which began in 1992.

Local authorities, are responsible for a wide range of essential services, such as education, housing, land use planning, social services and waste disposal. In addition, police (outside London) and fire services are funded and operated on a local basis. There is a total of 449 local authorities, consisting of 47 county councils, 333 district councils, 36 metropolitan borough councils and 33 London borough councils -- though work is on-going to rationalise the system. Together, they employ more than 1.5 million staff and are responsible for expenditure of over £44 billion.

Responsibility for the programme of results benchmarking lies with the Audit Commission. The Commission, formed in 1982 as a non-departmental public body sponsored by the Department of the Environment, with the Welsh Office, has the task of auditing local authority expenditure in England and Wales. From its establishment, it has conducted national studies designed to promote economy, efficiency and effectiveness in the provision of services by the bodies which it covers (these are considered in the section on process benchmarking, below). This work has now been complemented by the Commission playing the lead role in local authority results benchmarking.
Methodology

The Local Government Act 1992 for the first time required the Audit Commission to produce annual comparative indicators of local authority performance, including that of police and fire services. The resulting data is published annually.

The first year following the legislation was taken up with consultation between the Audit Commission and the bodies whose performance was to be covered. The process was complex and required sensitive handling, since local authorities are accountable to their own elected bodies, rather than to either the Audit Commission or Ministers. The agreed approach was for performance indicators to be defined for each area of activity. Each indicator was designed with the bodies whose performance it would measure, to ensure that the activity measured was appropriate and that the resources required for collection of the data were not excessive. The detailed methods by which performance was to be measured were published in 1993. Given the very wide range of activities undertaken and the number of areas selected for comparison, over 200 performance indicators were set. A selection of the indicators is provided in Annex C.

As part of the Citizen's Charter, councils and police forces had to publish in local newspapers the details of their performance against the indicators. This information, as well as an explanation of the system used for its measurement, was also supplied to the Audit Commission at the end of the year. The Audit Commission then collated the data and produced a commentary on the key activities to accompany its publication. The first set of data, covering the operational year 1993/94, was published in March 1995. The second set, covering 1994/95, was published in March 1996, thus starting the reporting of trends.

The Audit Commission's approach to the data has largely been to let the figures speak for themselves, although it supplies a commentary seeking to bring out key issues. The aim of the programme is to inform the public debate about the performance of public services. In publishing the information, the Commission has not, in most cases, attempted to define what constitutes good or bad service. In some cases this will be obvious but, in others, views will justifiably differ about whether or not a given level of performance is good. In addition, the Audit Commission has been at pains to ensure that the data are interpreted in a way that takes local circumstances into account, such as rural or urban communities.

Issues and problems

There was a fair degree of resistance to the programme in its initial stages. This was largely attributable to two factors. First, there was some perception that it represented a politically-motivated intrusion by central government into the affairs of local government. Second, many in local government predicted that publication of the data would have little impact and that the resources required for its collection would therefore be expended for little gain. These concerns proved to be unfounded.

The Audit Commission had undertaken research into the public's views and found that people valued the information being made available, believing that this would enhance public accountability. This was confirmed by the major press coverage which greeted publication -- over 2,000 pages in national and local newspapers and over 100 radio and television items were devoted to the subject. Concerns that reporting might be biased also proved unfounded -- research indicated that people tended to interpret the performance indicators as throwing a positive light on local government and were impressed at the range of services provided for the money spent. Press coverage was also more positive than negative. The result has been that there now exists a broad national and local political acceptance of the value of the performance indicators and of national publication.
The trend data published in 1996 revealed that those councils with the worst performance in the
previous year had improved significantly -- the five worst councils had improved performance by between
30 per cent and 60 per cent. However, a small number of councils still perform significantly below the
rest. Councils performing at or close to the average do not seem to have achieved improvements in
performance, yet comparison with similar councils, which are performing excellently for similar costs,
suggests that there is scope to improve.

Turning to police services, it was found that the performance of some services had deteriorated
despite significant increases in funding, while others with much smaller increases in funding improved
their performance significantly. Nevertheless, the Audit Commission noted that it can take time for
increases in funding to show through in improvements in services.

Use in decision making

For the future, the Commission aims to develop better measures of council and police
performance. The main thrust of its activity, however, will be to try to increase the impact of performance
indicators at a local level. The Audit Commission has no powers to require local authorities to use the
results benchmarking data as the basis for work to improve performance. It can work only by persuasion.
The Commission therefore intends to undertake a “roadshow” of seminars for councillors, police authority
members and officers, meeting them in their own council and police force areas and highlighting how the
indicators can help inform the process of performance review.

The Commission has also piloted an initiative to communicate directly with the public about the
performance of councils and police forces. In five main cities, it has installed a computerised version of
the performance indicators in the main library and will assess the level of interest from the public before
deciding whether to make this exhibition more widely available.

The Citizen's Charter Unit has also produced a CD-ROM that brings together performance
information for schools, hospitals, local authorities and emergency services. The programme uses a
computer map to help people find the information on their local services. Copies are being made available
free to schools, libraries and individuals.

International comparisons

A number of issues bear on whether the comparison could be extended to public services in other
countries. The benchmarking process requires information on the performance of specific activities to be
collected. It would require a significant international project to establish a common baseline, given the
varying approaches adopted by countries to the delivery of the services involved, but there would be useful
lessons to be learnt.

Process Benchmarking (1) -- Human Resource Management

Background

The July 1994 White Paper The Civil Service: Continuity and Change and the subsequent Command Paper The Civil Service: Taking Forward Continuity and Change acknowledged the very
substantial improvement that had been made in the efficiency of departmental and agency operations and
in the standards of service that they provide. Process benchmarking was advocated as one of the
management techniques which can help to improve further the efficiency and delivery of services. This section covers one of a number of initiatives being undertaken in central government to apply this technique to human resource management.

**Methodology**

In January 1996, ten central Government departments agreed to take part in a project, led by Development and Equal Opportunities Division, Cabinet Office (Office of Public Services), to explore the use of process benchmarking as a tool to improve human resource practices. The approach adopted was to form a consortium, where departments work together while at the same time carrying out separate projects. This provided a means by which participants could learn about process mapping and benchmarking, share their approaches, refine the techniques learned and develop sufficient skills to enable them to carry out benchmarking of other processes (not exclusively in the human resource area) in the future.

Following a tender exercise, consultants Coopers & Lybrand were contracted to work alongside the departments to help them along the benchmarking route. This firm has extensive experience in benchmarking (as well as management consultancy), including design of the European Best Practice Benchmarking Award.

It was recognised that measurements in themselves would not improve a process. Human resources has long been considered as an area of subjective and qualitative information not suited for “hard” measures. Benchmarking provides a way to harness the qualitative with the quantitative to provide measurable outcomes (see Annex D for the benchmarking route).

Individual projects were chosen by the participating departments covering the whole range of human resource functions -- such as personnel appraisal systems and measuring the impact of equal opportunities. From the outset it was emphasised that however small the scope of the project the issue should be something that was of critical importance to the organisation.

Although it is too early to comment on the outcomes of the individual projects, it is clear that the consortium will be a rich source of information -- on both the processes benchmarked and the benchmarking technique itself. The consortium report will include case studies of each project and there is already consensus that the collective learning and common issues contain excellent evidence that should be promulgated within and extended beyond the civil service.

**Issues and problems**

Different issues affected individual teams; however, there were some consistent key messages throughout -- some attributable to project management and others concerning the benchmarking process:

- benchmarking is only a vehicle to bring about the delivery of key changes and continuous development;
- those involved in benchmarking need to be personally motivated and have the skills to carry out specific roles - key skills include influencing, negotiation and problem solving techniques;
- a clear definition of the project must be agreed from the outset, setting the boundaries of the investigation;
– attempting to place hard measures to what was considered a “soft function area” was demanding - using process mapping and maturity profiles requires considerable effort;

– identifying key performance indicators for the organisation's own processes is essential to get a genuine insight;

– information needs to be contextualised for applicability within each organisation - do not adopt wholesale, be prepared to adapt.

**Use in decision making**

The findings of the project will be used to implement improvements in the organisations involved. It is intended to use the experience gained to inform future efficiency projects.

**International Comparisons**

Some of the projects lend themselves to international comparison. Although, there are no formal plans for the consortium to engage in international comparisons, one department will be visiting a parallel organisation overseas with this in mind. It is also expected that a number of the departments involved will wish to compete for the European Best Practice Benchmarking Award.

**Process Benchmarking (2) -- Value for Money Studies**

**Background**

An alternative approach to process benchmarking is represented by the use of efficiency scrutinies and value for money studies. Efficiency scrutinies were started by the first head of the Efficiency Unit in 1979, while both the National Audit Office (for central government) and the Audit Commission (for local government and, subsequently, the National Health Service) have employed value for money studies since the early 1980s. Since the techniques are similar, this section focuses on the work of the Audit Commission.

**Methodology**

Each year, the Audit Commission's two value for money units, focusing on local government and health and personal social services respectively, ask every local authority, police and fire service, National Health Service (NHS) trust and health authority what study areas they would find most beneficial. The responses are collated and used to inform the selection of ten to twelve projects to be taken forward, four or five of which will be in the local government area, the remainder covering health issues.

Each project involves a small team from the Audit Commission working with representatives of a group of service providers in the area under examination for a short period. The project team will seek to use comparative data from the participating agencies to identify best practice and bring out general areas for improvement. The final study report will give a picture of the current position of the service being considered and will make recommendations for improvement. These will be based both on best practice, and on options for alternative approaches. The resulting report provides a benchmark against which the performance of other service providers can be measured. Recent studies have investigated nursery
education for “under-fives”, local authority finance departments, hospital patient records, and NHS accident and emergency services.

Issues and problems

The Audit Commission has found that many of the bodies which it covers are keen to adopt the good practices highlighted in the Commission's national value-for-money reports before being subject to local audit. Such an approach can lead to strong local ownership of the change agenda. Building on its management handbooks, which help audited bodies implement locally appropriate solutions, the Commission is now experimenting with a new audit methodology called “study -- implementation -- audit”. This methodology has five key stages:

− early process benchmarking against the findings of a national value for money study, to provide a baseline of data against which an organisation's progress can be monitored;

− a diagnostic “snapshot” by the local auditor to ascertain the key local issues;

− the opportunity for the audited bodies to draw up their own “agenda for action” to address the main issues in the light of local circumstances and priorities;

− management handbooks, and possibly workshops, to support audited bodies when implementing their own solutions;

− a follow-up review by the auditor to establish how well the audited body has progressed against its own agenda for action.

Use in decision making

The Audit Commission's value for money approach to process benchmarking is well established and is becoming increasingly integrated into the decision making process by which these bodies review their performance and seek to identify areas for improvement and implement action plans. The Efficiency Unit's efficiency scrutinies and the National Audit Office's value for money studies play a similar role in central government.

International comparisons

Due to the specialised nature of the work, there are at present no plans to expand the programmes to include comparison with services in other countries.

Conclusions

The use of benchmarking techniques has been of increasing importance in the long-standing drive to improve the value for money offered by the public service in the UK, even though the term itself is of more recent adoption.

As in most other developed countries, the scope and scale of the public sector in the UK remains extensive and the bodies responsible for the delivery of services correspondingly large. The result has been
a host of initiatives, which have employed a range of benchmarking approaches, each focused on the particular circumstances of the area of work under consideration.

Process benchmarking is probably the longest established technique, dating back to the early 1980s, with the introduction of the Rayner (efficiency) scrutinies and the Audit Commission's value for money studies. This approach can be very labour intensive and depends on the existence of a sufficient number of bodies performing similar processes. In the right environment, process benchmarking has proved itself to be a powerful efficiency tool and has led to significant improvements in value for money. However, given the range of public sector activity this technique can be used only in a relatively small proportion of the whole.²

Results benchmarking has been adopted only since 1992, with the legal requirement being placed on the Audit Commission to publish data on local authority performance. This technique is highly effective where a number of organisations provide similar services. One of the benefits of results benchmarking in this circumstance is that there is no necessity to set a target for performance, thereby avoiding intrusion into the powers of elected local Councillors. The publication of data, identifying issues but without making judgements, has been shown to have a marked effect upon both public accountability and managerial attitudes to standards of performance.

Standards benchmarking is the most recent technique to have been used. The Charter Mark Award scheme has allowed high quality customer service to be recognised throughout the varied functions of the public services and has publicly set a standard in this area of activity for which organisations may aim. The Next Steps Agency benchmarking pilot seeks to go further, by measuring the performance of Agencies as a whole against the standards set by “best-in-class” businesses. The initial findings of this exercise indicate that this approach may be highly effective both at assessing the overall performance of organisations engaged in a wide range of activities and at providing a means for the identification and adoption of best practice from both within the civil service and beyond.

As the pressure becomes more intense for the public sector to seek ever more effective means by which to deliver high quality services to the public within increasingly limited resources, it can be expected that the use of the three main aspects of benchmarking will become a standard element in the efficiency techniques adopted by all sectors of the public sector in the UK.

² Some public services, particularly those that experience competition for example in the certification field, from international equivalent bodies following the opening up of the European Union market, have adopted the quality accreditation approach. Schemes such as the British Standard (BS5750), or the ISO 9000 equivalent provide a signal to prospective customers and an assurance about the internal processes of the organisation. While a case can be made for translating these standards, designed originally for manufacturing processes, to administrative and other service processes, central government in the UK has not generally pursued this route. This may reflect in part the strength of internal audit and inspection procedures and the newer focus on results and standards for the user, rather than on internal process itself. Thus, while a number of central Government laboratories and scientific institutions have sought (and achieved) quality accreditation, this has not seemed most appropriate for a widespread approach. However, in one respect, that of human resource management and staff development, the Government has promoted an accreditation scheme, Investors in People, which is being actively taken forward across Government as well as the wider public service and in the private sector.

25
ANNEX A

THE UK/EUROPEAN BUSINESS EXCELLENCE MODEL
ANNEX B

AGENCIES TAKING PART IN THE SELF-ASSESSMENT EXERCISE

<table>
<thead>
<tr>
<th>Agency</th>
<th>Staff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Army Base Repair Organisation</td>
<td>3 160</td>
</tr>
<tr>
<td>Army Base Storage &amp; Distribution Agency</td>
<td>5 175</td>
</tr>
<tr>
<td>Army Technical Support Agency</td>
<td>1 180</td>
</tr>
<tr>
<td>Civil Service College</td>
<td>250</td>
</tr>
<tr>
<td>Court Service</td>
<td>9 830</td>
</tr>
<tr>
<td>Defence Postal and Courier Service</td>
<td>560</td>
</tr>
<tr>
<td>Defence Secondary Care Agency</td>
<td>2 440</td>
</tr>
<tr>
<td>Driver &amp; Vehicle Licensing Agency</td>
<td>3 570</td>
</tr>
<tr>
<td>Driver &amp; Vehicle Licensing, Northern Ireland</td>
<td>215</td>
</tr>
<tr>
<td>Driving Standards Agency</td>
<td>1 585</td>
</tr>
<tr>
<td>Forensic Science Service</td>
<td>1 050</td>
</tr>
<tr>
<td>Government Purchasing Agency, Northern Ireland</td>
<td>60</td>
</tr>
<tr>
<td>HM Customs &amp; Excise</td>
<td>23 185</td>
</tr>
<tr>
<td>HM Land Registry</td>
<td>8 150</td>
</tr>
<tr>
<td>Inland Revenue</td>
<td>52 010</td>
</tr>
<tr>
<td>Intervention Board</td>
<td>855</td>
</tr>
<tr>
<td>Land Registers, Northern Ireland</td>
<td>215</td>
</tr>
<tr>
<td>Logistic Information Systems Agency</td>
<td>245</td>
</tr>
<tr>
<td>Medical Supplies Agency</td>
<td>260</td>
</tr>
<tr>
<td>NHS Estates</td>
<td>140</td>
</tr>
<tr>
<td>NHS Pensions Agency</td>
<td>430</td>
</tr>
<tr>
<td>National Weights &amp; Measures Laboratory</td>
<td>45</td>
</tr>
<tr>
<td>Patent Office</td>
<td>790</td>
</tr>
<tr>
<td>Public Record Office</td>
<td>445</td>
</tr>
<tr>
<td>Registers of Scotland</td>
<td>1 075</td>
</tr>
<tr>
<td>Social Security Benefits Agency</td>
<td>71 595</td>
</tr>
<tr>
<td>Social Security Child Support Agency</td>
<td>6 530</td>
</tr>
<tr>
<td>The Buying Agency</td>
<td>120</td>
</tr>
<tr>
<td>Veterinary Laboratories Agency</td>
<td>675</td>
</tr>
<tr>
<td>Wilton Park</td>
<td>30</td>
</tr>
</tbody>
</table>

**Total**                                         **195 870**

(Figures include some military personnel)
ANNEX C

SELECTION OF LOCAL AUTHORITY KEY PERFORMANCE INDICATORS

− The Audit Commission has established a series of more than 200 key performance indicators as a means to measure the activities of local authorities. The number is high, due to the wide range of the authorities' business. The Commission recognises that a balance must be struck between gathering sufficient data to allow an accurate assessment of performance to be made, while ensuring that the need to collect the data does not become an excessive burden.

− Given that the system has been in operation only since 1993, it is accepted that the indicators employed need to be adapted over time, in order to capture more accurately the key aspects of performance and also to minimise the resources required for their collection.

− The indicators listed below are only a small selection of the complete range, for which full details of local authority performance was published in April 1996.

<table>
<thead>
<tr>
<th>Area of Activity</th>
<th>Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Education</strong></td>
<td>% of 3 and 4 years old with an LA school place</td>
</tr>
<tr>
<td></td>
<td>Expenditure per primary school pupil</td>
</tr>
<tr>
<td></td>
<td>Expenditure per secondary school pupil</td>
</tr>
<tr>
<td></td>
<td>% of draft special educational needs statements prepared within 6 months</td>
</tr>
<tr>
<td><strong>Social Services</strong></td>
<td>% of elderly people over 75 helped to live at home</td>
</tr>
<tr>
<td></td>
<td>% of minor items of equipment to help people live at home provided within 3 weeks</td>
</tr>
<tr>
<td></td>
<td>% of adults going into residential care who were offered single rooms</td>
</tr>
<tr>
<td></td>
<td>% of children in LA care who are in foster homes</td>
</tr>
<tr>
<td></td>
<td>Number of children on the Child Protection Register per 1 000 children</td>
</tr>
<tr>
<td><strong>Libraries</strong></td>
<td>Number of books and other items issued by libraries per head of population</td>
</tr>
<tr>
<td><strong>Total Expenditure</strong></td>
<td>Total expenditure per head of population</td>
</tr>
<tr>
<td><strong>Council Housing</strong></td>
<td>Average time taken to re-let council dwelling</td>
</tr>
<tr>
<td></td>
<td>% of tenants owing more than 13 weeks' rent</td>
</tr>
<tr>
<td><strong>Recycling</strong></td>
<td>% of household waste recycled</td>
</tr>
<tr>
<td><strong>Planning Applications</strong></td>
<td>% of householder planning applications decided in 8 weeks</td>
</tr>
<tr>
<td><strong>Council Tax Benefit</strong></td>
<td>% of new council tax benefit claims processed in 14 days</td>
</tr>
</tbody>
</table>

*Table continued....*
<table>
<thead>
<tr>
<th>Area of Activity</th>
<th>Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Council Tax Collection</td>
<td>Council tax collected as a % of the total amount due</td>
</tr>
<tr>
<td>Inspecting Food Premises</td>
<td>Number of inspections of food premises as a % of the total inspections which should have been carried out</td>
</tr>
<tr>
<td>Quality of Complaints Systems</td>
<td>Quality of complaints procedures</td>
</tr>
<tr>
<td>Police 999 Calls and Emergencies</td>
<td>Performance in answering 999 calls</td>
</tr>
<tr>
<td></td>
<td>Performance in responding to emergencies</td>
</tr>
<tr>
<td>Crime and Detection</td>
<td>Number of recorded crimes per 1 000 population</td>
</tr>
<tr>
<td></td>
<td>% of all crimes cleared-up by primary means</td>
</tr>
<tr>
<td></td>
<td>% of violent crimes cleared-up by primary means</td>
</tr>
<tr>
<td></td>
<td>% of burglaries cleared-up by primary means</td>
</tr>
<tr>
<td></td>
<td>Number of crimes cleared-up by primary means per police officer</td>
</tr>
<tr>
<td>Police Resources</td>
<td>Number of police officers available for ordinary duty per 1 000 population</td>
</tr>
<tr>
<td></td>
<td>Proportion of police officers’ time spent in public</td>
</tr>
<tr>
<td></td>
<td>Expenditure on policing per head of population</td>
</tr>
<tr>
<td>Fire Service</td>
<td>% of fire calls at which “attendance standards” were met</td>
</tr>
<tr>
<td></td>
<td>Cost of the fire service per head of population</td>
</tr>
</tbody>
</table>
ANNEX D
CONSORTIUM PROJECT: BENCHMARKING HUMAN RESOURCES ACTIVITIES

Benchmarking Route

- setting up and developing a benchmarking team
- identifying human resources activities or processes which will benefit from benchmarking, and planning the project
- mapping the process and identifying areas for improvement
- identifying the best-in-class for the process
- arranging and organising site visits with the best-in-class to collect data
- analysing the data to produce an action plan aimed at delivering performance improvements
- winning support for and implementing the action plan
- monitoring results and developing a strategy of continuous improvement through benchmarking
ANNEX E

Agency Benchmarking Data

<table>
<thead>
<tr>
<th>Criterion</th>
<th>UK Quality Award Standard</th>
<th>Private Sector Average</th>
<th>Agency Average</th>
<th>Range of Agency Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leadership</td>
<td>63</td>
<td>60</td>
<td>35</td>
<td>48</td>
</tr>
<tr>
<td>Policy &amp; Strategy</td>
<td>68</td>
<td>40</td>
<td>36</td>
<td>57</td>
</tr>
<tr>
<td>People Management</td>
<td>66</td>
<td>50</td>
<td>36</td>
<td>50</td>
</tr>
<tr>
<td>Resources</td>
<td>68</td>
<td>60</td>
<td>41</td>
<td>52</td>
</tr>
<tr>
<td>Processes</td>
<td>75</td>
<td>50</td>
<td>35</td>
<td>53</td>
</tr>
<tr>
<td>Customer Satisfaction</td>
<td>60</td>
<td>30</td>
<td>38</td>
<td>61</td>
</tr>
<tr>
<td>People Satisfaction</td>
<td>66</td>
<td>40</td>
<td>22</td>
<td>45</td>
</tr>
<tr>
<td>Impact on Society</td>
<td>55</td>
<td>30</td>
<td>17</td>
<td>30</td>
</tr>
<tr>
<td>Business Results</td>
<td>86</td>
<td>50</td>
<td>46</td>
<td>63</td>
</tr>
</tbody>
</table>

This column provides Agency minimum scores used for blank base of columns in first chart.

Range of Agency Scores

Private/Public Comparison

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Private Sector Average</th>
<th>Agency Average</th>
<th>UK Quality Award Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leadership</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Policy &amp; Strategy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>People Management</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resources</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Processes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Customer Satisfaction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>People Satisfaction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact on Society</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business Results</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Public sector reform in Sweden has been significant at all levels of Government for more than a decade. Reform efforts are gaining momentum partly because of increasing difficulties in financing the public sector, coupled with mounting pressures caused by the globalisation of most aspects of social and economic life. A natural response has been the assimilation of new methods for improving performance. The benchmarking experiences summarised in this paper form part of this strategy.

**Dimensions in the Implementation of Benchmarking**

There are two dimensions that can be referred to in gaining a greater understanding of benchmarking experiences. The first distinguishes benchmarking as a process performed for either the evaluation or the continuous improvement of public activities. The second identifies the approach as being either results oriented or process oriented.

Granted the subjectivity and importance of “classifying” each benchmarking experience, it is important to keep in mind that benchmarking in all its forms contains a mixture of these four dimensions. These dimensions are to be understood as representing a rough or general interpretation of the prevalent direction identified in the four benchmarking cases presented in this paper.

- In the first case, that of the *Financial Management Rating of Government Agencies*, the direction points towards a process oriented benchmarking system that aims to initiate and consolidate continuous improvement in the financial management practices of 253 government agencies.

- The second case deals with the *International benchmarking of the Swedish Government Budget Process*. The principal direction here is undoubtedly process oriented yet with the objective of being evaluative in terms of offering an international perspective to a unique reform effort of the Government’s budget process.

- In the third case, that of an *International Comparison of Public Support for Families*, undertaken in eight OECD-countries, the emphasis is towards results oriented benchmarking for purposes of comprehensively evaluating the objectives and results of existing programs and services.
− In the fourth and last case, that of the Annual Benchmarking of Local Government Authorities, the approach is a results oriented one seeking to encourage and sustain continuous improvement in the activities performed by 284 local Government authorities.

**Diagram 1: Classifying Sweden’s experience with benchmarking in four different areas of public service**

---

**Financial Management Rating of Government Agencies in Sweden**

Substantial reforms delegating responsibility for financial and administrative matters, coupled with increasing demands for the improvement of accounting, follow-up and evaluation standards, have lead to major changes in the financial management practices of the Swedish State sector. One of these changes is the development and implementation of a rating system evaluating the financial management of government agencies.

**Purpose and principles**

In the early 1990’s, the National Audit Office was commissioned by Government to develop and implement a financial management rating (FM-rating) of all central agencies. The purpose of the FM-rating, is as an instrument for the Government in the regular follow-up of standards of financial management in agencies, and of results of changes in the division of responsibility between the Government and the agencies.

The FM-rating is composed of a two-letter rating for each agency:

---


---
The first position -- a so-called consolidated group rating -- represents an evaluation of financial management at the consolidated group level. It is meant to rate the relationship between the Government and the agencies.

The second position -- a so-called agency rating -- evaluates financial management at the agency level. It is intended to show how well the agencies meet general requirements in respect of financial management.

In both types of ratings the evaluation is performed on a scale of A to C. These are in turn combined in order to establish FM-ratings for each individual agency. The following groupings are used for determining a FM-rating:

- Group 1: AA and AB — Fully Satisfactory (FS)
- Group 2: AC, BA, BB and BC — Satisfactory (S)
- Group 3: CA, CB, and CC — Non-Satisfactory (NS)

In order to guarantee the quality of the rating and subsequent report to the Government, the National Audit Office, as an independent agency, is to determine each year the FM-rating of each agency. The National Audit Office is empowered to obtain any information considered necessary for the rating.

Financial management ratings for the fiscal year 1994/95

Beginning with the fiscal year 1994/95, it was decided to routinely implement the FM-rating for 253 agencies on an annual basis. The results of the rating for the fiscal year 1994/95 were only provided in aggregate form. The report to the Government covered groups or sectors of agencies and the identity of individual agencies was not specified. From the fiscal year 1995/96 and onwards FM-ratings of individual agencies will be identified and reported to the Government. The report for 1995/96 will be made in April 1997.

The following were the general results of FM-ratings for 1994/95:

Diagram 2a - FM-ratings for Fiscal year 1994/95

Diagram 2b - Consolidated Group and Agency Ratings

---

4 The report for 1995/96 will be made in April 1997.
The rating system

The consolidated group rating is based on the National Audit Office’s assessment of the agencies on the basis of a number of aspects that measure how the agency performs its financial management in the overall context of central government.

The following are examples of factors that are assessed:

− the audit report;
− overspending of appropriations;
− reports to the central government accounting system;
− cash management;
− reconciliation of payments;
− checks of closing balances and opening balances in the final accounts.

The agency rating is based on the answers the agency itself provides to a questionnaire that is then further processed, including examination by the National Audit Office of the agency. The questionnaire covers the financial management requirements that apply to all agencies. The questionnaire does not measure circumstances specific to the agency and is mostly of a formal character. It contains questions on items in the balance sheet and the income statement.

Examples of such items are:

− administration and valuation of outstanding claims;
− proportion of bad debt losses and interest paid on overdue payments;
− accounting for fixed assets.

Appropriation accounting is also covered. In addition, a number of general areas are measured to elucidate the agency’s methods of management, i.e.:

− procurement and sales of goods and services;
− integration with other information systems;
− business-like conditions for fee-financed activities;
− follow-up of earlier audit observations.

Evaluation of the FM-rating system

In 1996 the FM-rating system was evaluated by Stockholm University’s Department of Business Administration. Altogether 40 agencies were randomly selected and studied. Some interesting observations were:

− Three out of four agencies thought issues and areas covered by the FM-rating system were meaningful.
− Two out of three thought the FM-rating system measured the right aspects of financial management.

Sundberg Anna, Evaluation of The Financial Management Rating of Central Government Agencies for the Fiscal Year 1994/95, Department of Business Administration, Stockholm University, 1996.
More than half were positive to the idea the FM-rating of individual agencies being identified.

A similar number considered that the FM-rating system had not resulted in any changes in financial management. Two out of five took the opposite view.

Two out of five viewed the FM-rating system as being potentially valuable as a **benchmarking** device. More than half were negative to such an idea.

One out of two thought the FM-rating system had no value as a learning process. Close to an equal number shared an opposite view.

Two out of three were generally positive about being involved with the FM-rating system.

Thus, the evaluation of the FM-rating system’s first year of implementation conveys a divided picture. From a benchmarking perspective, both hazards and potentials can be identified, a principal one being: that close to two-fifths of the evaluated agencies expressed a **willingness to benchmark** their FM-practices in relation to the agencies that had earned top FM-ratings. Hence, next year’s individual FM-ratings stand a good chance of being used as envisaged.

---

**The International Benchmarking of the Swedish Budget Process**

**Comparing budgeting procedures and fiscal performance**

This case summarises the Swedish use of von Hagen’s (1992) model for analysing budgeting procedures and their relation to fiscal performance found in 12 EC countries for the period 1991-92. The model postulated the following two hypotheses:

- Differences in **long term “restrictions”** (macroeconomic objectives, transparency, budget periods, flexibility in implementation, as well as follow-up and monitoring procedures) are related to fiscal performance -- the so called “**long term hypothesis**”.

- Differences in **structural restrictions** (institutional arrangements restricting the actions of all participants and areas involved in the budgeting process) are related to fiscal performance - the so called “**structural hypothesis**”.

The testing of these hypotheses required the translation of qualitative descriptions of each country’s budget process to indicators (on a scale of 0-4), measuring 24 aspects involving the following areas of budgeting procedure:

- Dialogue within the Cabinet Office (restrictions, agenda-setting, initiative, negotiation);
- The Parliamentary Phase (limitations, proposal rights, veto rights, practices, financing);
- The Budget’s Level of Transparency (exclusions, uniformity, clarity, complexity, etc.);
- Flexibility in Implementation (authority, cash limits, transfers, control, schedules, etc.);
- Long Term Focus (macroeconomic goals, planning horizon, forecasts, commitment).

---

In order to construct indexes measuring each country’s overall level of restrictiveness in budgeting procedures, all indicators were subsequently added up to index-totals on a scale of 0-96 (higher values meaning greater restrictiveness). These indexes were then related to each country’s fiscal performance consisting of key economic indicators such as government fiscal deficit, net borrowing, etc.

More interestingly, the relationships between on the one side long term restrictiveness and structural restrictiveness, and fiscal performance on the other, were also analysed. This was done by altering the composition of the budget procedure areas included in each index.

All in all von Hagen drew the following general conclusions:

- The “long term hypothesis” could not be validated -- no relation could be established between fiscal performance and variations in long term restrictions

- The “structural hypothesis” was confirmed. Budgeting performance was found to be related to variations in structural restrictions. Regression analysis showed that as much as one-third of the fiscal deficit could be explained by structural restrictions, meaning that higher indicators for structural restrictions correlated with superior fiscal performance.

- Variations in the level of structural restrictiveness in each country’s initiating phases (the Cabinet Office Phase and the Parliamentary Phase) were the areas of budgeting procedure that were found to be most strongly related to fiscal performance.

The rating of the Swedish budget process

In late 1992, the Expert Group on Public Finance analysed the Government budget process following von Hagen’s model of analysis. The following conclusions were drawn:

On a scale of 0-80, Sweden obtained a rating of 20 for structural restrictiveness, to be compared with the EC country’s average of 40 and span of 20-70. In all four areas related to structural restrictions -- the Cabinet Office Phase, the Parliamentary Phase, the Degree of Budget Transparency, and the Flexibility in Implementation -- Sweden obtained lower indexes than EC averages. Long Term Focus was the only area where Sweden obtained a higher result. On the other hand, this area was concluded as being the one of least significance in terms of influencing fiscal performance.

On the basis of these results, the following changes were recommended:

- Strengthening structural restrictions over participants and activities belonging to the initiating phases of the budget process.

- The introduction of decision-making procedures enabling the fixing of an overall ceiling on spending during the initiating phases of the budget process.

- More stringent rules concerning MPs proposal rights.

- Improving parliamentary monitoring and auditing over policy implementation.
Major changes in the Swedish budget process

Over the past few years, Sweden’s Riksdag (Parliament) and Government have taken several decisions aimed at reforming and strengthening the budget process, an important element in the work of improving public finances.

One such decision is to introduce a framework budget model for handling budget proposals in the Riksdag. The model means that the Riksdag will first decide on the overall level of spending and then determine its distribution among various purposes.

A decision has also been taken that the Government’s budget proposals should come within a set expenditure ceiling for the public sector. This expenditure ceiling is defined in nominal terms, spans three years and covers the entire public sector (that is the state sector, excluding interest payments on government debt); social-insurance schemes; and the municipalities. For the state sector, the Government’s budget proposal is divided into expenditure areas. With the introduction of the ceiling, open-ended or draft appropriations as a form of appropriation have been abolished.

As far as the local government sector (the municipalities) is concerned the expenditure ceiling is somewhat different. It rests on a separate agreement between the government and the municipal sector.

Budget process and fiscal policy - past, present, and future

Along the same lines of the study conducted by the Expert Group on Public Finance, the Swedish Ministry of Finance recently evaluated the post-reform level of structural restrictiveness of the Swedish budget process. The results revealed a new picture conveyed by diagrams 3a and 3b below:

---

Diagram 3a - Pre-reform Structural Restrictiveness

Diagram 3b - Post-reform Structural Restrictiveness

---

On the whole, it is clear that both past and contemporary decisions aimed at reforming the Swedish budget process have been markedly influenced by the findings of the two earlier studies mentioned. However, it should be noted that the future effect on Sweden’s fiscal performance remains to be seen. However, a good measure of optimism may be warranted arising from the fact that budget process reform in Sweden is being strongly backed-up by a series of deep cuts in Government spending.

Comparing Public Support for Families: a Study of Eight OECD Member Countries

This is a summary of a study that the Expert Group of Public Finance commissioned from the Swedish Agency for Administrative Development. The purpose of the study was to compare public expenditures for transfer payments and services to families with children, existing in 1993 in Britain, Denmark, Finland, France, Germany, the Netherlands, Norway and Sweden. The study analyses and compares all forms of public support for families with children for the period beginning with a child’s time of birth to his or her completion of secondary education.

Findings and results

The study’s principal findings were: Sweden and Norway accounted for the highest levels of expenditure per child as well as for the implementation of the most varied systems of family support benefits and services.

---

* The study will be available in English early 1997.
Other important findings were:

- The highest expenditure levels for *child benefits (including tax subsidies)* were found in France and Norway; the lowest in Britain, Denmark and Sweden. Special priority was given in Denmark, Finland and Norway for child benefit systems to children before they start school. A similar status was given in the Netherlands, France, Sweden and Germany to child benefit systems directed to families with many children (these have since been reduced in both the Netherlands and Sweden). *Housing subsidies* were found to play an important role in France, Britain and Sweden. Child benefits together with tax subsidies and housing subsidies had the highest effects on *income redistribution* in Britain, France and Sweden (resulting from at least one type of subsidy being dependant on income).

- In the Nordic countries *parental leave expenditure* was substantially higher (Sweden displaying the highest level of expenditure).
– Denmark and Sweden accounted for greater expenditures for *child care* than the other countries. Practically no public expenditure was found to exist for this kind of purpose in both Britain and the Netherlands.

– Though being the area of highest expenditure in all countries, much less variation was observed in expenditure for *primary and secondary education*. The Nordic countries, however, showed substantially higher figures (Sweden and Norway the highest). Between the countries there were considerable differences concerning the *cost per pupil* in primary school. This was partly due to differences in the number of teachers employed.

**Method**

The study’s principal sources of information were written material from domestic and international organisations (OECD, Eurostat, the European Commission, EURYDICE, etc.), complemented by consultations with national and international advisers.

Comprehensive analysis was also carried out involving participation from “central advisers” for each country studied. This measure was taken for purposes of securing a proper evaluation and understanding of all information concerning each country. These “central advisers” were also allowed to read draft versions of the study.

Significant difficulties arose concerning the development and use of measures and indexes, mainly due to the effects of different tax systems, accounting methods, currency values and funding practices.

The principal measurement used for indexing each country’s overall level of commitment was that of *net public expenditure*. A net estimate was favoured for reasons of enabling the exclusion of both taxes paid on family benefits and of user fees for child-care services and education.

Each country’s *net expenses per child* was also indexed by dividing the net public expenditure for each benefit/service by the respective number of their *intended* users. These figures were then calculated for a *child’s first nineteen years of public support as if entirely occurring in 1993*. Further analysis was performed as these results were later summed up to totals revealing each country’s *combined expenditure* for all public forms of support existing in 1993 per child raised.

The study also made use of *purchasing power indexes* that related each country’s expenditure levels to the domestic value of local products and services.

A number of so called “*family types*” were developed to further analyse and compare the study’s principal findings in terms of *variations in income, number of supporting parents, as well as the number and age of children*. 

42
Policy issues in a Swedish context

The study discussed the following two principal policy issues concerning public programs and services for families in Sweden:

– Major differences were identified regarding the general nature of each country’s system for child benefits (including tax subsidies and housing subsidies). From a comparative stance, the Swedish system was found to be quite selective in nature (due to the fact that proportionately large housing subsidies were awarded on selective grounds). On the basis of this observation, the conventional perception held in Sweden concerning the general make-up of the domestic child benefit system was strongly questioned.

– Public expenditures for both subsidised services (child-care + primary and secondary education) and benefits (child benefits + tax subsidies + housing subsidies + parental leave), were found to be substantially higher in Sweden. Consequently, no general policy objectives assigning priority levels between subsidised services and benefits as two means of public support could be identified.

As a result, policy recommendations in the study centred on the idea that possibilities for reducing public expenditure for the support of families with children in Sweden could be identified. Reduced expenditure was seen as attainable by enhanced efficiency and/or by overall reductions in spending. Areas viewed as being suitable for reductions in expenditure were those of child-care, parental leave and education. This perceived suitability arose from the fact that even marginal reductions would translate to considerable savings in policy areas requiring such high rates of expenditure.

From a wider perspective, the making of well defined policy choices aiming at establishing priority levels and criteria for evaluating and comparing existing programs and services, was also highlighted as a desirable course of action.

The Annual Benchmarking of Local Government Authorities

This case deals with the annual benchmarking of the finances of Swedish local Government authorities that has been implemented since 1993. The benchmarking is carried out by The Swedish Association of Local Authorities in co-operation with the Central Bureau of Statistics. Its general purpose is that of elucidating the general financial results of local Government authorities.

The results are published in the form of a yearly report -- How Costly Is Your Local Government Authority -- containing information from the previous year’s annual accounts as reported by local Government authorities. Close to 140 different indicators for each local Government authority are developed and analysed each year.

Indicators reported

The annual report contains analysis of indicators describing all areas of the local Government economy. The indicators used are different to those used in other comparative studies that are often restricted to a specific area. For example, the National Agency of Education and the National Board of
Health and Welfare perform comparisons within their own areas. Both of these studies are considerably more detailed.

The basis of all types of indicators employed consists of statistics from the annual accounts of local Government authorities and of operational statistics, population data, and statistics on personnel.

The indicators can be grouped as follows:

- **Financial indicators** - data from income and balance statements furnished by local Government authorities. Analysis is also performed at the consolidated group level, i.e. the local Government authority including local authority-owned companies (there are about 1,300 such companies). Examples of financial indicators are different income/profit indicators, debt per inhabitant, solvency, etc.

- **Indicators for different welfare programs**. More comprehensive analysis is performed on education, child and elderly care. Examples of these types of indicators are: costs per inhabitant for an activity, schooling costs per student, care costs per pensioner, etc.

- **Indicators for transfer payments, personnel and employment**. Examples of these types of indicators are social security benefits per inhabitant, grant levels for recreational or cultural activities, number of employees per inhabitant, etc.

The analysis of indicators is reported either as: a percentage figure, as an index, as Swedish kronor per inhabitant, or as kronor per a given group targeted. In addition, analysis and comparisons are performed of different groups (size, etc.) of local Government authorities are performed.

Diagram 5 illustrates some of these indicators as averages for the entire local Government sector.

**Diagram 5 - Indicators for local Government authorities from the accounts for 1994**

<table>
<thead>
<tr>
<th>Revenue or expenditure item</th>
<th>1994</th>
<th>1994</th>
</tr>
</thead>
<tbody>
<tr>
<td>Government grants / resident</td>
<td>4500</td>
<td>8200</td>
</tr>
<tr>
<td>Tax revenues / resident</td>
<td>1200</td>
<td></td>
</tr>
<tr>
<td>Fees / resid. (excl. bus. act.)</td>
<td>2100</td>
<td>2100</td>
</tr>
<tr>
<td>Net expenditure / resident</td>
<td>3060</td>
<td>3060</td>
</tr>
<tr>
<td>Gross expenditure / resident</td>
<td>3690</td>
<td>3690</td>
</tr>
<tr>
<td>Labour costs / resident</td>
<td>3690</td>
<td>3690</td>
</tr>
<tr>
<td>Welfare benefits / resident</td>
<td>3200</td>
<td>3200</td>
</tr>
<tr>
<td>Job programs / resident</td>
<td>5300</td>
<td>5300</td>
</tr>
<tr>
<td>Elderly-care / user</td>
<td>6290</td>
<td>6290</td>
</tr>
<tr>
<td>Child-care / child</td>
<td>5360</td>
<td>5360</td>
</tr>
<tr>
<td>Secondschooling / student</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary schooling / student</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Variations between individual local Government authorities is substantial in many cases. For example, the expenditure for secondary schooling in the Stockholm metropolitan area spans from
46,000 kronor per student to 71,000. For primary schooling from 3,900 to 67,000 kronor per student, and for child-care from 42,000 to 69,000 kronor per child.

**The use of indicators**

The interest for following-up and evaluating local Government activities has increased significantly in recent years at both central and local Government.

Besides elucidating the annual state of local finances, a principal purpose is that of creating continued self-analysis concerning why differences exist between local Government authorities.

The need to follow-up the newly introduced Government grant system has further underscored the need for good indicators used in the financial practices of local Government authorities.

At the local Government level, this type of financial information is principally used when working with the budget. It is also increasingly being used independently as a basis for carrying out more in-depth comparative studies by one or more local Government authorities.

**General experience and problems**

The main problem with performance indicators of this character is the insufficient degree of comparability existing between different local Government authorities. The organisational and operational methods developed, together with the standards chosen for reporting financial and operational statistics, vary substantially between individual authorities. However, increased use has resulted in a continuing improvement of the base data used. Consequently, the continual task of improving the quality of data as well as that of the comparative process itself -- both in terms of the development and use of financial and operational statistics -- has been given the highest priority.

Beginning with the closing of annual accounts for the year 1995, a new system for classifying activities, costs and income has been introduced for the collection of financial data from local Government authorities. So as to improve the degree of comparability between local government authorities. Moreover, a Government commission is currently investigating proposals leading to the improvement of both the quality and comparability of financial information used in local Government. Addressing issues concerning the use of indicators in annual national follow-ups is a principal task being dealt with by the commission.

One important development aspect of the work with indicators -- being carried out by the Association of Local Authorities -- is that of finding the most relevant indicators that can be used at both the central and local levels of Government. An equally important task is that of devising models that can explain observed differences existing between individual or segments of local Government authorities.

**Swedish Benchmarking Experiences**

**Assumptions behind benchmarking**

There are several common traits concerning the motives driving the studies summarised in this paper. These can be grouped along the following lines:
– Perceiving the state of affairs and developments pertaining to the area(s) investigated as being generally unsatisfactory in terms of efficiency and effectiveness;

– Assuming that a benchmarking approach stands a good chance of identifying and analysing inefficiency in hitherto unscrutinised practices in the areas investigated;

– Sweden’s ambition to develop an internationally competitive public sector.

**What do the cases tell us about continuous improvement?**

In the two cases identified as aiming at continuous improvement -- the Financial Management Rating of Government Agencies and the Annual Benchmarking of Local Authorities -- the objective is explicitly stated as being the *recurrent benchmarking* of existing services and programs. Benchmarking is seen as a systematic and continuous process that once started, never ends.

Another similarity is the presence of benchmarking agents that are not directly responsible for the processes or results being benchmarked (RRV, Central Bureau of Statistics, Association of Local Authorities). Not surprisingly, this “outsider” effect can represent a challenge for the development of a benchmarking culture within the agencies concerned. This “ownership” aspect of the benchmarking process can substantially influence attitudes towards the concept of “learning from others”.

Diagram 6 illustrates a model for how agencies can implement such continuous benchmarking.

**Diagram 6: The Benchmarking process for continuous improvement**

The principles underlining this model are also evident in a handbook for benchmarking recently published by the National Audit Office.⁹

**Some insights on evaluation**

In the remaining two cases -- the international benchmarking of the budget process and the study in eight OECD Member countries comparing support for families -- the benchmarking performed was mainly evaluative in character. Hence, it centred on questioning the suitability of the systems or the results being achieved.

One important distinction between the two cases concerns their respective *initiating and implementing agents*. In the Budget Process case, the benchmarking was both initiated and performed by a government body found within the policy area being investigated. The contrary was true in the Support for Families case. This explains to a large extent why the benchmarking of the Budget Process had such a clear and immediate prescriptive outcome in comparison to the Family Support study’s more argumentative and suggestive conclusions. The findings of the latter case still forming the basis of present deliberation and public debate.

In both cases, whether centred on processes or results, the evaluative benchmarking’s effectiveness was determined in great part by the initial formulation of relevant policy issues that addressed the circumstances of the area(s) chosen for investigation. The complexity in terms of time, analytical skill and resources needed, explains why this type of benchmarking is normally performed occasionally or on an off time basis.

Diagram 7 shows a five-step procedure for how agencies can implement evaluation benchmarking that is partially deduced from the two cases.

**Diagram 7: The benchmarking process for evaluation**

This type of benchmarking plays an important role in *initiating* a public debate on the need for change., one that leads to the questioning of the established order without having to present a fully elaborated alternative policy, program or system. This perspective puts great emphasis on the right *timing* of evaluative benchmarking.

**The need for international benchmarking**

In recent time one fourth of the commissions appointed by the Swedish Government that have investigated different policy issues have made international comparisons or analysis of a benchmarking character. There are several motives encouraging this development:

- *Increased globalisation and international interdependency*. National policy frames are becoming too narrow for dealing with issues that are becoming more and more international in character (i.e. issues concerning finance, the environment, communications, energetic infrastructure, education, etc.).
– The absence of domestic comparators. International benchmarking enables or significantly widens the prospects for finding suitable comparators.

– The quest for competitive alternatives. In the face of change the international hunt for new ideas can stimulate positive change in areas containing obsolete and outmoded programs and systems.

– Widening the scope of debate and deliberation. By casting light on new arguments or aspects, international benchmarking can substantially widen the freedom of action for policy makers and implementers alike.

Implementing international benchmarking is neither simple nor inexpensive. There is a need to find ways to decrease the amount of time and resources usually needed. This may include making international benchmarking studies that are carried out in different OECD Member countries more easily available among all OECD members. There is also a need to analyse experiences on such issues as analytical approaches and the use of international benchmarking in the future development of public management.
The papers provided by the United Kingdom (UK) and Sweden illustrate the great variety of approaches to benchmarking, as well as the ongoing questions of how the benchmarks or criteria are determined and why they are considered relevant. Moreover, both papers demonstrate how useful -- if not necessary -- it is to classify the various approaches to benchmarking according to their scope of action and the methodology they employ.

The two papers differ with respect to the classification schemes. They demonstrate that the approach to classification may take various courses of action. Being faced with that very elemental difference between the papers, I had to decide whether or not to accept the different conceptual frameworks offered by the papers. That is, would I look at them from their differing perspectives, or would I try to develop my own comprehensive classification scheme in order to look at the papers in a co-ordinated way. I chose the second option.

In formulating my own classification framework, my first decision was to accept, in principle, the "quadrant classification" approach found in the Swedish paper. Quadrant classification means that two independent dimensions of classification are identified so that they can be presented graphically in the form of crossed axes constituting two pairs of alternatives. In the Swedish paper, the two dimensions are identified as process-oriented versus results-oriented, and continuous improvement versus evaluation. However, to enlarge the classification power sufficiently to accommodate both papers, it would be useful to go beyond the quadrant labels and assess the characteristics of each quadrant.

I accepted the distinction between results-oriented versus process-oriented as this distinction is also highlighted in the UK paper. By classifying benchmarking approaches as either results-oriented or process-oriented, or perhaps also results plus process-oriented, one can quite obviously understand a central characteristic of the activities being benchmarked.

The second dimension did not yield as readily to an easy characterisation. Here, the Swedish paper distinguishes between "evaluation" and "continuous improvement", while the UK paper emphasises "standards." To encompass both ideas, I decided to distinguish between two types of benchmarking. The first type of benchmarking uses preformulated normative standards as instruments of comparison and learning; whereas, the second type uses non-normative empirical data for that purpose. These empirical data might be ordered, compressed to indicators and processed in other ways in order to make them suitable for benchmarking.
Taken together, I now had a classification scheme consisting of two dimensions with two pairs of alternatives each. I thought this scheme might be quite useful for characterising -- at some essential points at least -- any benchmarking activity. Of course, to verify this assumption I tested it against both papers.

The UK experience

The UK paper reports on a pilot measurement of the "total activities" of 29 Next Steps Agencies. This measurement was done by means of the ASSESS RapidScore instrument of the British Quality Foundation (BQF) which consists of a broad range of normative standards that characterise well-performing organisations. On the background of my two-dimensional classification, it is important to note that this model includes process elements (so-called "enablers"), as well as result elements such as the objects of benchmarking.

The UK Citizen's Charter uses normative standards as means of assessing the client orientation of organisations. The task of the applicants to the Charter Mark is to prove that their activities are in line with a series of quality of service standards that have been set by the Charter Unit.

We find the opposite situation in looking at the Audit Commission's benchmarking of local authorities. Instead of normative standards, empirical data of achieved results (operationalised as "performance indicators") are collected and compared inter-organisationally. The intention is to "let the figures speak for themselves" without any attempt to define what constitutes good or bad service.

Finally, an even more dramatic contrast to the benchmarking of the Next Steps Agencies can be seen in the benchmarking of human resources management practices, as advocated in the White Paper "The Civil Service: Continuity and Change." This benchmarking exercise employs a combination of process related benchmarking measures. For example, it uses empirical data as means of evaluation and comparison. However, these data are not operationalised as "indicators," but rather collected and processed in a procedure called "process mapping". This process is not explained in the paper and therefore remains in need of further explanation.

The Swedish experience

Turning to the Swedish paper, one can find that the report on the financial management rating of government agencies in Sweden very clearly presents a combination of process benchmarking and results benchmarking, both using normative standards as instruments of comparison and evaluation. In methodological terms, this approach is, in principle, in full accordance with the UK pilot project for 29 Next Steps Agencies. Although in the UK case the "total quality approach" is not being pursued meaning that the scope of the standards included in the comparison is narrower and more specialised.

In sharp contrast to this combined, normative approach, Sweden's international comparative study of public support for families with children aims at collecting and comparing empirical data on the results of service activities. The same basic characteristic applies to the annual benchmarking of local government authorities. The only significant difference between the two approaches is that the scope of data collection is much broader for the local government authorities. In terms of the benchmarking methodology, both approaches are very close to that used in the UK benchmarking of local authorities, that, as I stated before, also concentrates on the collection and processing of results-related empirical facts.
Also in the Swedish case, the learning effect from benchmarking is expected from the approach to "let the figures speak for themselves".

Finally, the international benchmarking of the Swedish budget process also uses empirical data (without any normative meaning) as instruments of benchmarking. However, this approach takes one very significant step further by including not only results, but also process-related facts and by bringing these two groups of data into an "analytical" relationship with each other. Stated more concretely, this approach aims at explaining results through process-characteristics by exploring their statistical correlations.

Insights and Consequences

I now turn to the question of insights to be gained and consequences to be derived from benchmarking. I want to concentrate on those points that I myself regard as particularly important.

As stated earlier, the two papers represent a broad range of approaches to benchmarking. These approaches cover quite well -- maybe with some exceptions -- the main paths that are usually available. In view of this broad range, the question is raised whether there is any possibility to evaluate and rank the existing alternatives. Of course, the criteria for such an evaluation and ranking have to be selected very carefully beforehand.

On the background of the basic definition of benchmarking that I quoted in the beginning of the paper, it may go without saying that one main criterion will necessarily be the capacity and effectiveness of a benchmarking approach for fostering learning processes. With respect to the approaches under discussion, I want to put forward the thesis that they represent different learning strategies, as well as different probabilities and levels of learning effects to be expected.

Let us take, for example, the results-oriented comparison of empirical data that is found in the UK benchmarking of local authorities, the Swedish comparative study of public support for families with children, and the Swedish annual benchmarking of local government authorities. In all these cases, the learning effect to be gained rests almost exclusively on two factors: the impact of comparative statistics upon the general public and the actions of officials in the public sector who are expected to react to these statistics in an accountable way.

In this context, significant limits to the learning effect become visible. For example, the UK paper reports that only those councils that had shown the worst performance improved significantly after the publication of the relevant data; whereas, councils rated as performing at or close to the average had not improved even after some years, regardless of the actual scope of improvement. In other words, the strategy to "let the figures speak for themselves" produces the desired learning effect only if the comparisons discloses levels well below the norm. One can assume that the learning effect decreases to zero at an average level of performance, and speculate further that it may become negative above the norm.

Consequently, to get a higher level of learning capacity for effectiveness and optimisation, this type of benchmarking approach must be connected with additional elements of benchmarking or performance management. One strategy for improving the learning capacity might be to switch to a benchmarking approach that combines standards with a results orientation or, perhaps, combines standards with a results plus process orientation. This latter approach is used in several activities outlined in the papers, including the strategic concept behind the Citizen's Charter, the measurement of the "total activities" for Next Steps Agencies, and the Swedish financial management rating of government agencies. One could be inclined to expect higher learning effects, in particular, from the combination of standards
with the measurement of results plus process because this combination provides direct information on opportunities for improvement.

Another strategy for improving the learning capacity of data-oriented benchmarking might be to proceed in the direction of linking results-oriented data to process-oriented data in an analytical way. This is basically the strategy being pursued in Sweden's international comparison of budget processes. The inter-organisational comparison of empirical data on results will likely give rise to strong motivational factors (e.g. desire to have results compare favourably with other organisations, or to improve performance where there is a deficit compared to other organisations) that will increase the learning effect from benchmarking.

Before digging any deeper into the nature of the conditions for optimising the learning effect, it is necessary to consider additional criteria for evaluating existing alternative types of benchmarking. For example, the cost effectiveness of benchmarking has to be considered in addition to the learning effect. It seems to me that we cannot help but give weight to this criterion if we agree that benchmarking activities should cover as many comparable institutions or services as possible, and if we further agree that benchmarking activities should be institutionalised and repeated on a regular basis.

The introduction of a cost-effectiveness criterion leads to a ranking of benchmarking approaches that differs sharply from the ranking achieved when learning capacity is the leading criterion. For example, as the UK paper observes quite correctly, a process-oriented benchmarking approach that digs into a potentially voluminous store of data in order to analyse the causes of differences in results is labour intensive and costly. Furthermore, this approach can be used in only a relatively small proportion of the total activities of an organisation, because under this approach only activities and services of the same kind can be compared. Although there may be a high learning capacity attached to this type of benchmarking, it scores badly with respect to cost-effectiveness.

On the other hand, collecting comparative results-oriented data without process-oriented data scores badly in terms of the learning effect, as discussed earlier, but scores very highly with respect to cost effectiveness. This is because it can -- within certain limits -- be easily routinised and made compulsory, as has been demonstrated in the case of benchmarking of local authorities by the UK Audit Commission.

If we try to take both criteria in consideration, the evaluation of current benchmarking approaches runs into a problem of conflicting goals. However, in order to attain satisfactory results from benchmarking, the question of possibilities for reconciling the conflicting goals has to be addressed.

In my opinion, the first answer to this question may be that some very elementary infrastructure preconditions for benchmarking should be guaranteed in the public sector. Stated more concretely, there should be an institutionalised practice to select and operate performance indicators that concentrates on results. These indicators should be agreed upon across the public sector, as far as is possible. On the basis of a continuous stream of information gained by the indicators, various supplementary activities could be developed. A growing number of indicators could be used as a basis for formulating normative service standards. Likewise, various groups, or "rings", of benchmarking partners could be established to use information on performance levels as a starting point for making improvements and benefiting from the lessons offered by the "best-in-class" examples. The results of such projects could also be published and used by others outside the group. I should like to mention briefly that in Germany we are proceeding quite well on this track. Stated with some optimism, we are approaching a common catalogue of performance indicators that can be used by most local governments. Furthermore, a growing number of benchmarking groups made up of similar organisations have been developing. They proceed from the comparison and
critical discussion of results-oriented data on the identification, analysis, and adoption of best-in-class (or best-in-group) practices.

The motivational power and learning capacity offered by benchmarking could be increased by using best-in-class performance management practices as a basis for formulating operational standards that can be used by all organisations as benchmarks for self-assessment and improvements. In this context, award competitions that follow the total quality management (TQM) principles can be very helpful as reinforcers of learning-oriented attitudes. In Germany, we have created the Speyer Quality Award for the public sector, that ran successfully in 1992 and in 1994, and that now (in its third round) is also available for Austrian and Swiss candidates. Award competitions mainly score very high in terms of learning capacity, but there are problems concerning cost-effectiveness. In this regard, the main problem may be the rather limited number of public organisations willing to take the risk of coming forward as candidates. However, the impact of this problems can be reduced by organising reward competitions as "bench-learning" events, and by providing for a "medium probability" for all candidates to be among the winners.

I would like to end by mentioning that we ourselves are not close friends of the ISO-certificate programme. Its seems to us in Germany that this certificate is too much oriented toward formalisation and documentation, and too little oriented toward the promotion of learning and innovation. As a result, its respective capacity and effect for improving performance management is rather low.
EVALUATION OF PUBLIC MANAGEMENT REFORMS IN FINLAND:
FROM AD HOC STUDIES TO A PROGRAMMATIC APPROACH

by Katju Holkeri and Hilkka Summa,
Ministry of Finance, Finland

Introduction

Governments pursue public management reform policies to achieve certain goals. Such goals usually reflect the wish to improve some aspect of the performance of public bodies or the public sector as a whole. Given the emphasis on performance and result assessment in recent reforms in many member countries, it is natural that the reform programmes themselves should also be evaluated. In the past few years, the Public Management Committee (PUMA) has often stressed in its publications the importance of evaluation to public management reforms. It has pointed out that evaluation is important for the successful implementation and for the sustainability of reforms.

Discussions in PUMA meetings and publications have noted that little systematic evaluation of public management reforms has been done in the Member countries (Surveys on public management developments 1990 and 1993). This may be related to the difficulties of evaluating broad and multi-faceted programmes, but also to the fact that public sector reforms may be considered as irreversible, once-and-for-all undertakings based on political choices, the success of which will be measured by the support that the elected decision makers enjoy. The implementation of a reform may be considered as a result in itself, rather than an activity to be evaluated.

In Finland, discussions were already underway in the early 1990s about the importance of evaluation to the proper development and implementation of reform policies. In the year 1992-1993, the Public Management Department of the Ministry of Finance (that in Finland is the unit responsible for policy advice and implementation of public management reforms) carried out an international comparison of public administration developments (the Kanske project). The project found that no thorough evaluation of the recent Finnish public management reforms had been made and suggested in its final report (World's Best Public Sector, 1993) that the Ministry of Finance should start evaluating the reform policies that had been carried out in Finland since the mid 1980's. However, it took some years before a programme was started for this purpose.

This paper will describe an on-going evaluation programme launched by the Finnish Ministry of Finance in 1995 to assess the results of the public management policies pursued in Finland during the years 1987-1995 (i.e. during the terms of the two previous governments). As the work is still in process, the paper will concentrate on the goals, methodology and steering of the programme, as well as on the expected use of its results. At this point in time, only preliminary remarks can be made about the actual findings of the programme.
From Ad Hoc Studies to a Programmatic Evaluation

Major public sector reforms were carried out in Finland during the years 1987 to 1995. Both the conservative-social democratic coalition government (1987-91) and the subsequent centre-conservative coalition (1991-1995) gave serious attention in their programmes to the need to modernise the structures and functioning of the public sector as a whole. The present social-democrat led coalition government (1995) has continued this reform policy, but with a slightly different agenda.

During the years 1987-95 a number of major reforms were launched and fully implemented. Substantial changes occurred in financial management, personnel management and the structure of the state sector. The most significant parts of the reform programme were:

− adoption of a results-oriented budgeting and management system;
− devolution of decision making authority on personnel management to the agency level;
− reform of the system of state grants to municipalities;
− corporatisation of a large number of state bodies, converting them into either state enterprises or companies;
− structural reforms such as mergers and restructuring of state agencies and reorganisation of the provincial level of government.

For most of these reforms, the planning and implementation phases took several years. For example, the results-oriented budgeting and management system was adopted in an incremental exploratory process between the years 1989 and 1995. Early 1995 was a time when several major reform efforts had been completed. Also at that time a new government came into office and started planning its program for the coming electoral period. Thus, the autumn of 1995 was a logical time for launching a comprehensive evaluation of the reform policies carried out to that point. In addition, for the first time the Finnish government had a minister who specialised in administrative affairs and public management reform. In the previous governments, administration and public management reform policy had been an additional task within the portfolio of a minister for some other sector.

Doubts on the utility of evaluation

The first proposals for a major evaluation undertaking did not meet only with positive responses either among civil servants or political decision makers. It was pointed out that evaluation results would not be usable because the decisions on the reforms had already been made and would not be reversed based on evaluation findings.

There were also concerns that an evaluation questioning the merits of a reform program would harm the proper implementation of the policies. Other issues that delayed the launching of the evaluation programme were:

− the financial crisis that focused politicians' attention elsewhere;
− loss of interest because the government was coming to the end of its term, and there was concern that this long term effort would not produce results in time for the ministers to respond;

− changes in the leadership of the department in charge of planning and directing the programme, that delayed the decision making.

**Phase One: relying on spontaneous interest of academics**

The Public Management Department has a tradition of liaising with university departments in the fields of public administration and management. Therefore, it was natural to look to the university sector for potential evaluators and interest in evaluation studies.

There was some spontaneous interest by academics and research institutes in the public sector reforms. However, this did not develop into high-profile or comprehensive studies. In Finland, public management reforms traditionally had not been high on the political agenda, nor had they been the subject of broad public debates. There is no well established school or body of academic research on topics related to public sector reforms, and the number of university departments or other research institutes working in the fields of public management or administration is small, and none of them specialise in programme or policy evaluation.

In May 1994, the Public Management Department arranged a seminar with the aim of arousing the interest of academics and research institutes in studying and evaluating public sector reforms and changes, and to help civil servants involved in the implementation of reforms to network with researchers in the same field. Altogether about 200 civil servants and researchers from universities and other research institutes participated in the seminar. The results of the working groups were published in a volume entitled, “The Unknown Reform” (Tutkimaton Uudistus, 1994), that gave an overview of the need to evaluate public sector reforms, and of the academic research that was going on in the field at that time.

A number of interesting, evaluatory studies on the changes in the Finnish public sector were published during the first years of the 1990’s (e.g. Naschold, 1995, Kasvio & al., 1994), without any influence from the Ministry of Finance. Some studies were commissioned by the municipal sector and others by other ministries. The Ministry of Finance also commissioned individual studies on selected aspects of the reform policies, mainly on the basis of proposals put forward by individual researchers and research groups (e.g. Public Services go on the Market 1994, Salminen & Niskanen 1996, Puoskari 1996).

**Phase Two: towards a programmatic approach**

It soon became clear that despite an increasing amount of ad hoc interest generating research on reforms, this was not the way to get a comprehensive and systematic view of what had been achieved and what still needed attention in the reform process.

In late 1995, the Department decided to take a more programmatic approach to the evaluation. The on-going and recently finished evaluation projects that had been commissioned by the Department were gathered under one “umbrella project”. A steering group was set up to co-ordinate and steer the evaluation programme with a view to building up a comprehensive programme with specified goals, a fixed schedule for carrying out the programme, and a plan for using and publishing the results. To help the planning of the programme a pre-study was commissioned from the Political Science Department of
the University of Helsinki that set out and weighed the main theoretical and methodological options for a comprehensive evaluation of the reforms (Temmes, 1994).

**Why evaluate?**

There may be different objectives for an evaluation of reforms. Assessment of the success or failure of certain reforms may be one motivation. For example, a new government may wish to have an assessment of the achievements of a previous government before making its own policy choices. Another motivation may be to assess the performance of reformers (policy advisors or implementors) themselves. A third may be the need for a basis for better informed policy choices. Often evaluation is the only way to get an overall picture of the complex effects, including possible unwanted side-effects, of a reform programme.

In the Finnish case, the last of these possible motivations was the most important one. The overall objectives of the evaluation programme are more forward looking than backward looking. The key objectives are to:

- provide a basis for the government to decide on the points of emphasis in its programme during the rest of its term;
- get an overall view of how the public sector has changed during the past decade -- i.e. does it operate differently, does it cost less, is it more efficient, how do the various interested parties view the changes;
- compare the Finnish reform policies with other countries to establish a benchmark on where Finland stands internationally in the quality of its public sector.

Indeed, more important than judging the success or failure of the previous governments' work is the learning effect: where are we now and what should be done next. It was also noted that the “evaluation culture” in the Finnish administration was still rather thin. There is no strong tradition of using information produced by evaluations for more informed decision making. Thus a secondary motivation for a systematic evaluation was to learn about evaluation itself: to gain experience in planning, steering and using evaluation in the policy process.

The Finnish evaluation effort is characterised by a pragmatic approach. No sophisticated attempts are made to solve all the methodological and theoretical problems that are bound to arise in complex evaluation studies. The aim was to develop a programme that could be completed in a reasonable amount of time and within a reasonable budget, and to produce usable and relevant results based on data, methods and expert judgement that were sufficiently credible for decision makers to rely on them. An evaluation striving for a very high level of sophistication and scientific depth might be useless if it came too late for the government to take action based upon its results.

**Methodological Issues**

Evaluation of broad reform programmes is never a simple task. Policy choices in reform programmes are based on beliefs and expectations about what improvements the changes will produce. These expectations may be based either on successful experiences from the past or from other countries, ideologically grounded views of the proper role, size or modes of operation of the public sector, on more theoretical grounds concerning the good functioning of organisations or, more often, on a combination of
these motivations. There already exists a substantial amount of documented experience and a growing body of evaluative academic literature against which the validity of these views can be checked. Yet, the effects of individual reforms, let alone broader reform programmes, are usually complex and dependent on multiple factors and contextual developments. Thus, an objective assessment of the success or failure of reforms needs to be based on multi-faceted information on the conditions of implementation and other factors influencing the perceived results.

There are a number of methodological problems that have to be solved in each evaluation project. These concern, the criteria to be used in the evaluation, the time period to be considered, and the attribution problems in distinguishing the effects of reforms from the effects of other factors.

Criteria

The first choice to be made when planning an evaluation is to decide whether the evaluation should focus only on the achievement of the official objectives of the programme, or on the effects of the programme in a more general way. The former type is called a goal-bound evaluation, and the latter type is called goal-free evaluation. A goal-free evaluation attempts to study the impacts of the programme irrespective of whether these have been directly stated as its concrete objectives, taking account of, e.g., possible unwanted side effects or the relevance of the programme objectives considering changes in the environment since the programme was launched. Thus a goal-free evaluation may introduce a wider set of criteria.

When evaluating a government programme the most usual choice is a goal-bound evaluation, particularly when the civil service is the evaluator or commissions the evaluation. However, the problem that arises is that very often policy objectives are vague, ambiguous or even contradictory.

Reform policies are often a multi-faceted combination of over-all policy decisions, individual reform projects and detailed changes in regulations. This is particularly true in the Finnish case. The reform programme carried out from 1987 to 1995 is more a series of individual reforms than one coherent programme with a single set of objectives. Moreover, the reform objectives are scattered across several policy documents, government decisions-in-principle, committee proposals and motivations for legislation. As a result, the coherence of the reform programme seemed to be a target for the evaluation rather than a starting point for it.

The Finnish evaluation programme can best be described as a goal-free evaluation, based on a set of criteria that attempt to cover those issues that will be most relevant for future decision-making on the reform process. The evaluation criteria are as follows:

- the effects of the reforms from the citizens' point of view -- i.e. the citizens' views on the level and quality of public service;
- the governability of the public sector from the point of view of political decision makers, including its adaptability to changing conditions;
- the efficiency and effectiveness of the functioning of public bodies and the public sector as a whole;
- the working atmosphere, quality of leadership and motivation of staff in government organisations.
The idea of the evaluation programme is to assess to what extent, if any, positive changes have happened in the workings of government organisations when judged against these impact-criteria.

In addition, the following criteria related to the process of the reform programme as a whole were included:

- the quality of the implementation process of the reforms; and
- the coherence of the programme and the relevance of its objectives from today's perspective.

Studies of individual components and aspects of the reforms were included in the programme so that sufficient coverage of all these perspectives could be obtained with the programme as a whole. However, the on-going programme is a compromise between the need for collecting data for proper operationalisation of all the criteria, and time and budget constraints. Original new data is collected only in some of the studies, whereas some aspects are covered by secondary material or by experts’ assessments gathered by interviews.

**Time period**

When planning an evaluation of the effects of a reform a problem to be solved is the time period during which the results of the reform are expected to mature. An evaluation of the effects of a change or a reform means that a comparison has to be made between the situation before and after the reform. In the case of a public management reform programme defining both these points in time may be problematic. The reform programme -- as is largely the case in Finland -- may consist of a series of changes and individually decided reform projects, that have a cumulative effect over a long period of time. In addition, the reforms may have multiple and ambitious objectives that cannot be expected to materialise overnight. In the Finnish case the adoption of the results-oriented budgeting and management system within the state administration was motivated by, among other things, a need to change the “administrative culture” towards a more result-oriented and service minded approach. The result of an evaluation on whether or not this objective has been achieved will depend on when the measurements of the effects of the reform are made. Assumptions will have to be made about the lapse of time needed for the “final” effects to mature.

In Finland, there was a natural basis for the choice of the time period. The two governments that were in office in the years 1987 to 1995 had public management reforms high on their agenda, and several national level reform projects were started and fully implemented during this period. Therefore, the main focus of the evaluation programme was confined to this period. Yet, it is obvious that it is still too early to give a “final judgement” of the effects of the reforms. On the other hand, the usability of the results for the present government's decision-making on how to continue the reform process requires that evaluation results be made available as early as possible.

To get a fuller picture of the effects of reforms, some individual studies within the evaluation programme will look at changes over a longer period of time. For these studies, only interim results will be included in the evaluation programme. For example, a study on how the adoption of results-oriented management and budgeting is influencing leadership culture and staff motivation will be carried out as a follow-up survey with repeated measurements over a three years period, but only the results from the first round of questionnaires will be included in the final report of the evaluation programme.
**Attribution problems**

Methodological difficulties often arise in defining and getting evidence for assumed causal relationships. This is particularly true in the evaluation of public management reform programmes where multi-faceted “packages” include a range of different decisions, recommendations, new legislation and other measures. For example, on what grounds could observed changes in the efficiency of public organisations be attributable to specific reforms? Pollitt (1995) identifies methodological difficulties typical in the evaluation of public management reforms:

- Even if a particular effect or impact can be identified, how can the evaluator decide which facet or component of the reform programme produced it?
- More often than not other changes in the political/administrative environment are going on at the same time as the reform programme (e.g. budget cuts -- that was the case in Finland). How can the evaluator know whether a particular effect (such as an increase in productivity) is generated by the reform programme, or by the simultaneous shifts in other variables?
- Even if a particular effect could confidently be attributed to a particular reform, what would be the appropriate comparator for the new state of affairs? The level of performance of the public organisations before the reforms may not be an appropriate comparator because in many instances their performance may have been improving (changing) anyway, for reasons unrelated to the reforms or other recent environmental changes (e.g. technological developments).

It seems obvious that conclusive evidence of cause-effect relationships can seldom be established in an evaluation of public management reform policies. This understanding was taken as a starting point in the Finnish evaluation programme, that is a compromise between methodological purity and the need for usable results as a basis for the “next steps” to be taken. The focus of the Finnish programme is more on mapping how the reformed public sector organisations and mechanisms are working now, than on establishing in-depth evidence on the detailed effects of individual aspects of the reforms. However, in one of the individual studies original data will be collected with the aim of testing the validity of a theoretical model on the effects of the results-oriented management and budgeting reform on leadership culture and staff motivation.

**Carrying Out the Evaluation**

The evaluation programme consists of 12 complementary studies, chosen so as to cover the major reforms carried out during the time period, and the different evaluative perspectives defined by the four impact criteria and the two process criteria mentioned earlier.

**Steering the programme and reporting to decision-makers**

As mentioned earlier, the programme was started in an “incremental” way, with the Ministry of Finance commissioning some individual evaluative studies on particular reforms. To get a more comprehensive approach, a steering group was set up in December 1995 to plan and co-ordinate the evaluation programme. The timeframe given for the work was approximately 15 months, with the final report expected in March 1997. The first task of the steering group was to plan and initiate new studies so that all the relevant aspects of the reforms and the different evaluation criteria would be covered. The steering group had also carried out a fairly intensive review of the individual studies during the course of
the studies. This was considered necessary prior to initiating new studies in order to raise the relevant questions and to help the researchers get the factual information on the aims and contents of the reforms.

The steering group has seven members, with a chair and secretary from the Public Management Department (the authors of this paper), and the other members representing the Prime Minister's Office, the Ministry of Interior, the Budget Department of the Ministry of Finance, the Office of Parliamentary Auditors and the State Audit Office. All of these organisations have been involved in the planning and implementation of the reforms and/or are potential users of the results of the evaluation. This composition, along with the fairly intensive co-operation between the evaluators and the steering group, will raise questions on the objectivity and neutrality of the results of the evaluation. However, it seems difficult to organise a broad evaluation programme so that this problem could be totally avoided. Any evaluator will be dependent on the information supplied by those involved in the reforms. In the Finnish case, a balance was sought between objectivity (commissioning outside academic evaluators who will publish the individual studies in their own names) and relevance and usability of the results (directing the researchers to relevant questions and making sure that they get all the information they need). In addition, the international comparison that is included in the programme (the FINREF project discussed below) will bring in an outsider's (i.e. non-Finnish) view.

The ministerial working group also identified the need to strengthen the use of evaluation in policy planning and implementation in general in the Finnish administration. This gave the steering group another task besides planning and co-ordinating the evaluation programme itself -- i.e. to find means and make proposals for strengthening the evaluation culture in the Finnish administration. The activities under way in relation to this task are briefly described later in the paper.

**Choosing the researchers**

It was clear from the start that the Public Management Department would not do the actual research work for the evaluations, but that the Department would commission the evaluation studies from outside producers.

When choosing the researchers, emphasis was placed on:

− sufficient variety of research institutes and other producers;
− sufficient variety of disciplines;
− the researchers' expertise in the structure and functions of the Finnish administration.

The researchers for the 12 individual studies come from different Finnish universities and research centres, except for the international comparison that has been commissioned from a British university. The choice of researchers for each study was made individually with an attempt not to concentrate on too few producers. This is a concern in a small country such as Finland, where there is no actual market for these kinds of expert services. This is also the main reason why a “tailored” approach rather than competitive tendering was chosen for selecting the research institutes to be involved in the studies.

Another issue in choosing the producers of the individual studies was to ensure wide enough coverage of the different disciplines and research traditions. The programme now includes researchers from administrative and political sciences, national economists, psychologists and business economists.
Finally, the Ministry of Finance also recognised that the researchers needed to have a thorough knowledge of the Finnish administration and its administrative systems.

The Twelve Evaluation Studies

The whole programme consists of 12 studies commissioned from different research institutes. The goals, methods and time tables for each of these are briefly described below. Preliminary results are also mentioned where available. One of the twelve studies is an exploratory report on how to design the evaluation of the Finnish public management reforms (Temmes, 1994). This study is not included in the following descriptions.

The table provided in the Annex describes how the studies cover the different evaluation criteria and the different reforms implemented during the study period.

The studies fall into two major groups:

Group A: where the focus of the study is mainly on one particular reform.

Group B: where the focus of the study is on the reform programme as a whole or on its effects in general.

Group A studies include:

1) Public services go on the market -- Experiences and views of new public enterprises and companies.

Goals of the study: To gather the experiences and evaluations on the process and effects of reorganising agencies into public enterprises or companies in the Finnish public sector, that occurred in the early 1990s. The aim was to get a general picture of the situation as viewed by the management, personnel and owners of the new public enterprises and companies. The inquiry focuses on changes in the operating environment, organisation and management, and development of businesses in the corporatised bodies.

Methods: A questionnaire was sent to the representatives of management, personnel and owners of 20 state and municipally owned enterprises and companies, (supplemented by interviews in 10 enterprises) asking their views on changes in their respective organisations, measured on the following 7 dimensions: relations with the environment (e.g. competition), leadership and decision-making, financial management, personnel management, quality of service, client feedback and its use, relations to the owners. Each of these dimensions was operationalised with a set of specific questions.

Timetable: The study was initiated before the evaluation programme was set up and results were published in 1994 as Public Services Go on the Market.

Results: There had been an increase in demand for the services of the reformed organisations and a rise in their number of customers. Their profitability had generally improved. They have increased their marketing activities, developed new products and discontinued unprofitable services. There are increased efforts to systematically gather customer feedback. Personnel management has developed in a more flexible direction, e.g. with an increase in fixed-term contracts. However, remuneration systems have not changed remarkably. There has been a reduction in control from the owners. Overall, the evaluation is positive regarding views on
profitability and quality of service. However, representatives of personnel criticise the manner in which the changes have been carried out and management practices.

2) Implementation of reform policies in the Ministry of Trade and Industry and in the Ministry of Interior.

Goals: The study analyses whether and how the goals of public management reforms, particularly the introduction of results-oriented management, have been met and been their effects have been in these two ministries. These ministries were chosen because they were known to be managerially “fairly advanced” because they had started their reform processes early on and because they welcomed the opportunity to take part in this study. The reforms studied included results-oriented management, commercialisation of subordinate units, projects for improving service, and reorganisation in the central, district and local administrations of these ministries. The questions investigated were:

– to what extent the targets of reform policy have been met;
– factors that have helped to achieve the targets;
– restraints that have hindered achieving the targets;
– possible deficiencies in the reform policy;
– unwanted side-effects of the reforms;
– positive developments resulting from factors other than the reform policy.

Methods: Interviews with 16 experts. For the Ministry of Interior an inquiry was also sent to 10 provincial governments and to the 20 recently created regional federations of municipalities. In addition, planning and monitoring documents of the ministries were used, as well as earlier studies from this area.

Timetable: Results were published in 1995 (Ahonen et Virta, 1995).

Results: In the Ministry of Interior the reforms have been implemented effectively. The main effect has been a general change towards emphasising clients’ views and needs. Criticisms were made of the quality of negotiations and performance consultations in results-oriented management and of how streamlining of provincial governments was implemented. In the Ministry of Trade and Industry, especially in larger units, many of the targets for the reform policy had been achieved (e.g. establishing the target setting and follow-up systems, increased focus on results). However, the relationships between the Ministry and its subordinate organisations remained somewhat unclear.

3) Effect of the results-oriented management system on the organisational culture of government agencies.

Goals: To investigate how and to what extent management practices and organisational cultures in government bodies change as a result of the introduction of results-oriented management and budgeting. Questions studies included, e.g. to what extent are management practices, staff attitudes and commitment to work actually changing at the unit level, and how far are the observed changes related to elements of the results-oriented management system, such as flexibility in resource use, explicit target setting and follow-up on performance.

Methods: Changes in organisational culture and management practices will be measured over a period of three years. Data will be obtained from a questionnaire in each year 1995 to 1998 sent
to all employees in 17 government organisations. The sample of organisations includes large and small units, both latecomers and pioneers in results-oriented management, and units both with easily measurable tasks and tasks that are difficult to measure. Cultural change and management practices are operationalised with a large set of questions covering relationships between superiors and staff, details of management practices, the nature of commitment to work and the organisation, and attitudes towards performance.

**Timetable:** The study will publish a report each year with a final summarising report in 1998. The results from the first year will be included in the summary report of the evaluation programme.

4) Effects of the reforms on regional and provincial administrations.

**Goals:** To study how the reforms of provincial and regional administrations have influenced the structures, steering mechanisms and personnel of the public bodies at these administrative levels. The reforms covered are:

– transfer of the regional development functions from the state provincial governments to the regional federations;
– transfer of environmental functions of the provincial level from provincial governments to separate environment centres;
– streamlining provincial governments;
– enlarging the responsibilities of district administrations, e.g. the police and the customs services.

**Methods:** The evaluation is based on the six criteria provided below. The first five criteria will be measured by a questionnaire to the mayors of all the municipalities (who are considered the main customers of the bodies in regional administrations), and data for the sixth criterion (costs) will be gathered from documents. In the questionnaire, the five criteria were operationalised through a set of specific questions. Additional data to get a full picture of the changes will be gathered by interviews of managers in the organisations involved, and later another questionnaire will be sent to other customers (besides municipalities) of the organisations studied.

– how far the “bottom-up principle” (that was an objective of the reforms) has actually been realised in regional development functions;
– the quality of the co-operation between the bodies involved in regional development functions;
– the effectiveness of operations of the bodies influenced by the reforms;
– the quality of service;
– the clarity of administration; and
– the financial point of view (focusing mainly on costs).

**Timetable:** Final report by the end of February 1997.
5) Citizen’s views on public services.

**Goals:** To survey the views of citizens on public service organisations today and to assess whether and how the changes in public management have influenced these views. The study is based on comparisons between different sectors of public services and between public and private services, and a comparison between the situation in 1987 and 1996.

**Methods:** This is a repeat of a similar survey made in 1987. The respondents (a sample of 2,000 citizens from the general population) were asked for their views on 12 government services, 11 municipal services and 13 private services, based on a set of 11 characteristics from which they were to choose two that they thought described the service in question. The characteristics are evaluative descriptions some of which are positive (flexible, trustworthy, service-minded, etc.) and some negative (old-fashioned, inefficient, etc.). The questionnaire will be supplemented by interviews with the managers of the government organisations included in the study to get possible explanations for the changes in the results between the 1987 and the present study.

**Timetable:** Final report by the end of 1996.

**Results:** The preliminary results indicate that there has been a positive change in how citizens view those government service organisations that were commercialised in the early 1990s. For example, the vehicle inspection centre and the post offices are now seen to be more service-minded, flexible and efficient than at the time of the previous survey.

6) The financial results of commercialised agencies.

**Goals:** To assess the financial effects of the conversion of government organisations into public enterprises or companies.

**Methods:** Using a sample of five reformed organisations to gather time-series data on key business indicators (turnover, gross margin, costs, prices, production volume, investments, etc.) over a period of approximately 10 years (i.e. from when the unit was first an agency, then a state enterprise and then a company).

**Timetable:** Final report by the end of 1996.

*Group B* studies include:

7) Views of senior civil servants on public management reforms.

**Goals:** To get a varied and comprehensive picture of the opinions of the senior civil servants on the public sector reforms of the 1990s.

**Methods:** The study was based on interviews with 28 senior civil servants in state administration and in the Association of Finnish Local Authorities. The interviewees represented different administrative fields, ministries and agencies, as well as different levels of government. Permanent state secretaries of many ministries were included in the interviewees.

**Timetable:** The study was published in 1995 (Temmes and Kiviniemi, 1995).
Results: The attitudes towards reform policies were generally positive. This positive picture included the direction and results of the reforms. Criticisms fell into three categories. First, there was criticism about the reforms being generalised and not sufficiently sensitive to variations between different tasks and administrative fields. Secondly, the reform programme’s formal goals diverged from its operational goals and there were contradictory elements, especially in the reforms concerning the structures of central and regional administrations. A third line of criticism related to the timing of the reforms and the scarce resources given to their implementation.

8) Politicians’ views on the success of reforms.

Goals: From 1987 to 1995 there was a ministerial committee in the Finnish Government in charge of guiding and monitoring the public management reforms. This study aimed at getting the committee members’ assessment on the reform achievements.

Methods: The ten longest serving members of this committee (all former ministers) were interviewed. They represented eight ministries and six different parties. The method was a semi-structured theme interview, covering the ministers’ experiences and assessments of the reforms and their views on the future changes and development needs. In addition, assessments of the working processes of the ministerial committee were sought.

Timetable: The report was published in 1996 (Kiviniemi, 1996).

Results: The interviews showed that the main lines of the reform policies were generally still accepted by the ministers. Based on this study and the study of the opinions of the senior civil servants (see earlier discussion), there was not a big difference between the opinions of these two groups on reform policy. The views of the politicians seemed to be more principle-oriented and general, whereas the civil servants emphasised pragmatism and concentrated more on details. The ministers saw the main results of the reform policies as: 1) streamlining the administration (smaller size, lower costs), 2) increasing efficiency, 3) enhancing the service orientation of the administration. The ministers accepted managerialism as the correct reform policy, but considered that the implementation of reforms had been uneven, with the same line of reforms producing positive results in some fields and negative effects in others. Political decision-makers put more emphasis than civil servants on the citizens’ point of view, a global view of the political and administrative system, and the needs and functioning of the local (municipal) level.

9) Financial and economic effects on the public management reforms.

Goals: To obtain an overall assessment of the effects of public management reforms on the micro (productivity and unit cost-efficiency) and macro (the control and management of public expenditure as a whole) levels. The reforms studied are results-oriented management, net budgeting, corporatisation and the reform of state grants to municipalities. The study will examine the costs, numbers of staff and productivity in the activities affected by these reforms, both on the unit level and on the level of the public sector as a whole.

Methods: The analysis consists of:

– a summary of what effects these reforms could be expected to have, assessed on the basis of relevant scientific (microeconomic) literature and previous research, focusing particularly on how the reforms will change the incentive structure faced by public organisations.
– an empirical analysis of the effects of reforms, based on overall statistical data on unit costs, productivity development and numbers of personnel in different sectors of public activities. This analysis does not aim at conclusive evidence of the exact cause-effect relationships between observed changes and the reforms, but rather at providing indications of the likely effects.

– an analysis of productivity potential and factors influencing productivity at the individual organisation level. This will be a comparative summary of recent empirical productivity studies on a number of state organisations conducted with the DEA (Data Envelopment Analysis) method (the policy force, land surveying offices, employment offices and tax offices).

**Timetable:** Final report by the end of 1996.

10) An international comparison of the Finnish public management reforms (*FINREF* project).

**Goals:** The strategy behind this study is to compare the implementation of public management reforms in Finland with those in four other countries -- the United Kingdom, New Zealand, the Netherlands and Denmark. The aim of this comparison is to see if there are useful lessons for Finland. The main method is to search the primary and secondary literature, and on this basis to produce accounts of the direction of change in each country, and then to validate these accounts through interviews with a small number of key actors and experts.

**Methods:** Literature analysis and interviews, both structured by a set of key questions concerning implementation processes. The questions were agreed between the Ministry of Finance and the FINREF project team at an early stage. They are too numerous to reproduce here but include such issues as:

– The scope of the reforms?
– How coherent was the logic of the reforms?
– The balance/mix between different types of reform, e.g. privatisation, market-type mechanisms, decentralisation, customer-responsiveness, etc.?
– Administrative mechanisms to steer the reforms?
– How and how far were the reforms evaluated, and with what results?

It is already clear that considerable attention needs to be paid to the political and administrative systems within which management reforms are implemented. A particular reform may “fit” one country well but be difficult or impossible, or simply irrelevant, in another. Because of this, great care has to be taken in “borrowing” reforms internationally. Nevertheless, the FINREF study will attempt to identify specific areas where there may still be scope for further reform in Finland, or for making existing reforms more effective.

In addition to the five country comparison, the FINREF project includes three institutional case studies. Each of these is an Anglo-Finnish paired comparison. One compares reform in the UK Treasury and the Finnish Ministry of Finance, the second is of management changes in primary and secondary education, and the third concerns reforms in postal and telecommunications services. These case studies aim to further illuminate the broad themes emerging from the country
comparisons and to permit a more in-depth examination of the detailed processes of implementation.

Overall the FINREF project, whilst utilising output and outcome data on those relatively rare occasions when they are available, will focus mainly on process criteria, i.e. assessments of whether management reform programmes have been implemented in speedy, coherent and effective ways, and on the extent to that they have been systematically evaluated.

**Timetable:** Final report by the end of February 1997.


**Goals:** The final report will summarise the findings of other studies and present an overall evaluation of the reform programme. It will also summarise the changes that have occurred in the Finnish public sector from 1987 to 1995, focusing on structures, costs, numbers of personnel and steering mechanisms of the administration as a whole. The study will also draw conclusions about how the reform policy can be continued and further developed in the coming years. The main contents of the final report will be:

- starting points of the evaluation programme;
- comparisons of the state of the public sector in 1987 with the state in 1995;
- analysis of the reform processes and parties involved in it;
- conclusions about the success of the implementation of the reform policies;
- a summary of the central findings on the impacts of the reforms.

**Methods:** The work on the final report is being carried out in close co-operation with the researchers conducting the other studies in the evaluation programme, as well as with the steering group of the evaluation programme.

**Timetable:** Final report will be ready by the end of March 1997.

**Use of the Results**

The success of any evaluation depends, in the final analysis, on whether its results will provide usable input to decision making on future programmes or on the revision of on-going ones. Timeliness and relevance of the topics to be evaluated, as well as the criteria used, are important prerequisites for a usable evaluation. Attention also has to be paid to how and when the results are published and forwarded to the decision-makers. Publishing a report is seldom sufficient to make sure that the results of evaluations will reach decision-makers when the information is most needed and most likely to have an influence on the formation of new programmes.

When planning the Finnish evaluation programme several measures were taken to ensure that its results will be relevant and inform the discussions on future reform needs:

- The steering group consists of persons who are directly involved in the planning of new reforms. As the group keeps itself informed of the studies as they progress, the group members will be able to take account of the findings as soon as they have any relevance for the rest of their work.
– The steering group reports to the ministerial working group that ensures that the expectations of the decision-makers will be known and that findings will be directly discussed with them.

– The time span of the evaluation programme was set with an aim to have the final results early enough to be used as input when planning the next stages of reforms.

– The results will be widely distributed, supported by seminars and other measures to stimulate public debate on the direction of change in the public sector.

– In addition to the individual studies, two overall reports evaluating the reform policies as a whole will be produced. A summary report will put special emphasis on recommendations and findings that are particularly relevant for the future work. The report on the international comparison will give a benchmark for where Finland stands in relation to other countries that have taken comparable reform strategies, and thus also help in finding neglected areas and possibilities to improve the process and implementation of the reforms.

The reform programme as a whole will produce an overall picture of the transformation of the Finnish public administration since the latter half of the 1980’s, and an informed view of the role the reforms have played in this transformation. The most important use of the results of the programme will be to identify further initiatives for moving towards a “government that works better and costs less”. This final result of the programme will be jointly produced by the evaluators and the steering group.

**Strengthening the Evaluation Culture in Finland**

The steering group of the evaluation programme has as its second task finding ways of strengthening the evaluation culture in the Finnish administration.

To start this work, a mapping was made of the current state of evaluation practices in the Finnish central government. This survey showed that there is great interest in these questions generally in the administration. Even if in some sectors the use of evaluations is well established and methodologically advanced (such as development aid and higher education), in most ministries evaluation is mainly done and used on an *ad hoc* basis. The overall picture was that ministries needed to strengthen their use and knowledge of evaluation. Evaluations of individual organisations, such as government agencies and research institutes, are becoming routine in some sectors. For example, the Ministry of Trade and Industry has stated an evaluation programme covering all its subordinate agencies. However, at the same time, policy and programme evaluations are still rare. The survey showed that the importance of evaluations in political steering of the administration is quite weak. It also showed that the use of evaluation results, as well as the competence of civil servants in planning and commissioning evaluations, are areas that would need to be strengthened.

The Ministry of Finance, together with the Finnish Institute of Public Management, has planned a series of training courses for civil servants responsible for evaluations in their ministries. The training aims to disseminate best practices in how to design and commission evaluations, and in how to make best use of the results. Another measure is the preparation of a handbook/guide on good practices in evaluation. As background work for this handbook, the Public Management Department made a comparative survey on similar handbooks published recently by other central governments of OECD Member countries.
The question of how to use evaluations and whether there is a need for statutory requirements, e.g. for obliging ministries to develop evaluation programmes for their policy areas, is also discussed in a Ministry of Finance project for further developing the results oriented management system. The idea is that the ministries would report on their evaluation programme and the results of evaluations in their annual reports.

Conclusions

At this stage it is still too early to present any final conclusions from the evaluation programme. However, some general remarks can be made on the basis of the experiences thus far.

First, it is obvious that a broad evaluation programme will need broad support both from political decision makers and civil servants to ensure that the effort will be worthwhile. The idea of increasing and improving evaluation is generally supported, but when it comes to starting any major evaluation on specific subjects doubts often arise. Ex-post evaluations of projects where decisions have already been made may be considered useless. Also, fears may arise that an evaluation may reveal failures or problems that would be difficult to correct. Criticism from external evaluation may be considered useless -- even if it would not be opposed as such -- if there is no channel to use the results in the preparation of new programmes. Very often evaluations are dealt with as an activity separate from programme planning and implementation that reduces support for and interest in evaluation.

Secondly, the need to evaluate is taken into account in planning the reforms is a prerequisite for good evaluation. In the Finnish programme this was not the case. The issue was recognised, but not considered as a key factor for starting the reform programme. Including a strategy for evaluation in the planning and decision making phase of future reforms will sharpen the definition of objectives. The steering group will propose in its final report that in the future these principles should be recognised and be standard procedure.

A third set of observations concerns balancing the needs for objectivity with sufficient expert knowledge. Objectivity and neutrality are vital for an evaluation to be trustworthy and useful. When evaluating public management reforms, two types of expert knowledge are needed: skills to use evaluation methods and techniques, and knowledge on the specific structures, procedures and culture of the administration in question. In a small country it may be difficult to find evaluators who would at the same time be sufficiently competent in evaluation methods, well acquainted with the administrative procedures and independent of the civil service and political decision makers whose work is to be evaluated. In the Finnish programme a balance has been sought through the co-operation between the steering group that consists of “insiders” and evaluators who are independent “outsiders”.

An important element in providing objectivity to the Finnish evaluation programme is also the international comparison (the FINREF project) commissioned from a British research institute specialising in public policy evaluation. While the project is still under way, it has already proved to be an excellent way to get a “bird’s-eye-view” of the Finnish reform policy and to provide fresh views on the alternative reform paths. However, an international comparison is not a simple task if it is to produce new information beyond the general trends that are already widely known. A special problem in international evaluations is the deep structural and cultural differences in how public management is discussed and understood between countries. It is time consuming for the evaluators to gain in-depth understanding of these differences. This problem is multiplied where there is a language barrier. Both these problems mean that giving adequate support to an international team of evaluators requires resources in the “home base”. In the Finnish case there has been a team assisting the evaluators, the work of which has included translation.
and producing surveys and summary reports on those aspects of the reforms that were included in the comparison. In practice, international evaluations based on field work with new original data collection will seldom be possible (except within same language areas).

A fourth conclusion at this interim phase of the programme is that it seems evaluations of public management reform programmes present specific problems that may not be as prominent in policy evaluation in general. Public management reform programmes tend to be comprised of a series of policy statements, decisions and projects that accumulate over a long period of time, rather than launched as a coherent programmes with a simple set of objectives. It may be hard to define the boundaries of a programme leaving it unclear what should be included in an evaluation. Another problem is that the objectives for public sector reforms may be stated in terms of vague “improvement” needs instead of specific measurable goals. In addition, the “maturation” time for the effects of reforms may be long and hard to define, that means that evaluation should be done as a continuous follow-up rather than as a once-and-for-all project. On the other hand, decision-makers may require firm and clear results from evaluations at a stage that is actually too early for any final judgements. Therefore, a balance needs to be found between practical needs and scientific and methodological sophistication.
### ANNEX

The Finnish Evaluation Programme

<table>
<thead>
<tr>
<th>Evaluation study (The numbers of the studies refer to those in the text)</th>
<th>Focus of the study</th>
<th>IMPACT CRITERIA</th>
<th>PROCESS CRITERIA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Evaluation study</strong></td>
<td><strong>Citizens' view of public service</strong></td>
<td><strong>Steering possibilities from the point of decision-makers</strong></td>
<td><strong>Efficiency/Effectiveness.</strong></td>
</tr>
<tr>
<td>1. Public Services Go on the Market</td>
<td>Corporatisation</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2. Implementation of reform policies in two ministries</td>
<td>Mainly results-oriented management</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>3. Effect of results-oriented management to organisational culture</td>
<td>Results-oriented management</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Reforms of provincial and regional administration</td>
<td>Structural reforms</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>5. Views of citizens</td>
<td>Mainly corporatisation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Financial success of state-owned enterprises</td>
<td>Corporatisation</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>7. Views of civil servants</td>
<td>Reform programme</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>8. Views of ministers</td>
<td>Reform programme</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>9. Financial results</td>
<td>Reform programme</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>10. International comparison</td>
<td>Reform programme</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>11. Final report</td>
<td>Reform programme</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| 73 |
REFERENCES

AHONEN, P. and VIRTA, S. (1995), Julkisen sektorin uudistukset ja niiden vaikutukset -- sisäasianministeriön ja kauppa -- ja teollisuusministeriön hallinnonalat (Public sector reforms and their effects in the Ministry of Interior and in the Ministry of Trade and Industry), Ministry of Finance, Helsinki. (not available in English)


KASVIO, A. and al (1995), Uudistamisen voimavarat -- Tutkimus kunnallisen palvelutuotannon tuloksellisuuden ja työelämän laadun kehittämisestä (The resources for transformation -- A study on the effectiveness of municipal service production and of developing the quality of working life), University of Tampere. (not available in English)

KIVINIEMI, M. (1996), Poliittisten päätöksentekijöiden käsityksiä hallinnon uudistamisesta (Views of political decision-makers on the public management reforms), University of Helsinki. (not available in English)


Public Services Go on the Market (1994), (a shortened English version of the report “Julkiset palvelut menevät markkinoille”), Ministry of Finance, Helsinki.

PUOSKARI, Pentti (1996), Transformation of the Public Sector, Ministry of Finance, Helsinki.

SALMINEN, A. and J. NISKANEN (1996), Markkinoiden ehdoilla (On the Terms of the Market), Ministry of Finance, Helsinki. (not available in English)


TEMMES, Markku (1994), Hallinto puntarissa, (Administration Weighed), Ministry of Finance, Helsinki. (not available in English)

TEMMES, M. and M. KIVINIEMI (1995), Muutoksen mahdollisuudet -- Arviointea ja näkemyksiä julkisen sektorin ja hallinnon kehittämisestä 1990 luvulla (Possibilities for Change -- Assessments and Views of Developing the Public Sector and Public Management in the 1990s), Ministry of Finance, Helsinki. (not available in English)


Tutkimaton uudistus (1994), (The Unknown Reform), Ministry of Finance, Helsinki. (not available in English)
COMMENTARY

by Danièle Lamarque
Cour des Comptes, France

Evaluation Of Public Management Reforms In Finland

Introduction

The public management reforms implemented in many OECD Member countries must be evaluated in order to assess their effect on the quality, efficiency and effectiveness of public services. But the evaluation of modernisation experiments gives rise to specific methodological problems (Trosa, 1992; Pollitt, 1995) and the reference material is still not sufficiently conclusive, as little has been done thus far to evaluate public management reforms systematically.

The value of the Finnish report is that it uses a practical example to illustrate a systematic approach to evaluation and to highlight the main methodological difficulties of evaluating public management reforms. The evaluation that is under way in Finland was launched in the autumn of 1995. It is based on a number of studies that have already been conducted, and is aimed at assessing the results of the public management policies which were implemented in Finland between 1987-1995, reforms which comply with the usual characteristics of the New Public Management (adoption of a results-oriented budgeting and management system, devolution of personnel management, structural reforms of public bodies and local authorities, conversion of government bodies into public or private enterprises, changes in the financial relationship between central government and the local authorities).

The six main methodological issues highlighted by the report are as follows:

- evaluation objectives;
- definition of criteria and preparation of the reference framework;
- time period, or the duration of the programmes concerned;
- causal relationship: are the effects observed due to management reforms or to other factors?
- follow-up to evaluation.
Evaluation objectives

The main objective of evaluation in Finland is to improve public management. The focus is firmly on the future. The three objectives are:

- to guide the action of the new government (elected in 1995);
- improve knowledge of how the public sector works, and particularly of the changes which have occurred over the last ten years;
- provide reference information based on comparisons with other countries in order to show where Finland stands internationally in terms of the quality of its public sector.

The evaluation is intended to have a dual learning effect: the aims are to improve knowledge within the administration of how the public service works, and to strengthen the public sector’s evaluation culture.

The sponsor for the evaluation project is the Public Management Department, which is part of the Ministry of Finance, and is responsible for implementing public management reforms. It also chairs the evaluation steering group. The close involvement of the sponsor and its central place in the public management reform system are a guarantee that the evaluation conclusions will be taken into account in the public management system. Therefore, in this approach, evaluation is quite closely integrated with management activity.

A precise definition of evaluation objectives is a central issue, as was recalled in 1996 by the French Scientific Council of Evaluation in its *Petit Guide de l’Évaluation des Politiques Publiques* (handbook for the evaluation of public policies):

> Considering the context that has led to a call for evaluation, the expectations expressed by the sponsor and the phase in the life cycle of the policy concerned by the evaluation, the aim is to identify the type of practical results and spin-off which can be expected from the operation, as well as its possible consequences for the different stakeholders. These consequences may be quite different depending on whether the evaluation is limited in scope (technical adjustment of the procedure, restatement of certain objectives), whether it may result in substantial change in the allocation of budget resources, lead to the winding-up of a programme or, on the contrary, to its generalisation, etc.

It is essential to take into account the sponsor's expectations, particularly when a reform is being evaluated (Trosa, 1992). In this respect Finland’s approach gives rise to two types of question:

- Is there a risk of a contradiction between the evaluation goal and the focus on pragmatism of those in charge of the evaluation who want to keep it relatively short (15 months) and limit its costs so that results can be rapidly obtained as a basis for decisions? The difficulty is partly resolved by using studies which have already been launched, or even published, and by issuing intermediate results for those studies that require more time (such as the analysis of the effect of management reforms on organisational culture in government agencies).

- Even if the evaluation includes studies on the public’s expectations, it is mainly based on the viewpoints of those responsible for management (politicians or civil servants). In addition to
its sponsors, should the evaluation also take into account its other potential users and the various stakeholders, i.e. those whose interests may be affected by the follow-up to the evaluation (e.g. handbook issued by the French Scientific Council of Evaluation in 1996)?

Definition of the criteria and preparation of the reference framework

A basic issue is to determine with what criteria should the results observed be compared? On the basis of what standards can it be said that the policy is working? The Finnish example shows, once again, that the objectives of the policy evaluated cannot be the only legitimate framework for the evaluation. This is different from the relatively narrow definition adopted in France in the Decree of 22 January 1990 for the creation of the interdepartmental system of evaluating public policies:

_Evaluating a policy is to see whether the legal, administrative and financial resources used produce the effects expected of this policy and achieve the objectives assigned to it,_

The French example of the evaluation of government computer systems (in 1992) showed that evaluation may affect the grading of objectives and promote latent objectives. This evaluation, which initially focused on efficiency (at the outset the sponsor was interested in the return on investment as measured by effectiveness in terms of productivity gains compared with resource allocations), moved towards a more complex concept by taking into account the quality of the service provided for the public, staff working conditions, procedure security, etc.

The Finnish approach avoids the usual pitfalls in evaluations of modernisation programmes, where the criteria are mainly derived from official programme objectives, such as improvement in efficiency, management devolution, improvement in the quality of service, etc. (Pollitt, 1995). It defines four impact criteria which relate to public expectations concerning the level and quality of public services, the freedom of movement for policy decision-makers in public service management, the efficiency and effectiveness of the public sector, staff working conditions, and two process criteria, one of which relates to the way in which the reforms are implemented, and the other to the coherence of the programme and the relevance of its objectives. However, the difficulty of preparing a reference framework and establishing stable benchmarks to measure the progress attributable to management reforms (Pollitt, 1995; Trosa, 1992), inherent in this type of evaluation, still has to be settled.

The time period, or the duration of the programmes concerned

The common difficulty in evaluations of public management reforms is that these reforms are usually too recent for all their effects to be perceived (Trosa, 1992; Pollitt, 1995). Fortunately this is not always the case. In France, the decentralisation programme, that has profoundly modified the public management system by developing the local authorities’ responsibilities and autonomy, dates back to 1982. Therefore, by 1995 the Cour des Comptes (French Audit Office) had sufficient time to evaluate its effects in the fields of secondary education and welfare.
The reforms analysed in the Finnish evaluation were launched in the years 1987-1995, which is a sufficiently long period for the effects of at least the earliest reforms to be identified. This period also corresponds to political markers (the two governments which were in power in that period) consistent with the political drive necessary for the launch and implementation of management reforms.

**The causal relationship: are the effects observed attributable to the management reforms or to other factors?**

The issue of causal relationships is seen as one of the most difficult in an evaluation exercise, particularly in the case of public management reforms which are generally multi-faceted, relate to a diversified context and are usually accompanied by other important changes affecting the operation of public services. Although this difficulty is not ignored, a strict scientific approach is sacrificed to pragmatism in the Finnish evaluation, which is meant to be “a compromise between methodological purity and the need for usable results”. However, in one study original data are collected in order to test the validity of a theoretical model that analysed the effects of the new management methods on leadership culture and staff motivation.

**Evaluation procedure**

The operation is supervised by a steering group with seven members comprising representatives from the main ministries concerned, Parliament and the Audit Office. The sponsor, the Public Management Department, provides the Chairman and Secretary. This administrative structure, limited to the parties mainly concerned, shows that the evaluation is closely connected with actual management. This is a guarantee that the evaluation findings will be used. It should be recalled that this is one of the main objectives of the evaluation.

However, this system involves a twofold risk:

- It does not guarantee the objectivity and neutrality of the final conclusions, even if transparency is partly ensured by publishing specific studies under the responsibility of the researchers who have conducted them.

- To a large extent the process is closed, at least in the steering phase, to outside participants, whereas the evaluation body should make it possible to compare the viewpoints of those directly responsible for the policy being implemented with those of external experts, stakeholders, etc. (Lamarque, 1993, French Scientific Council of Evaluation, 1996). The French Scientific Council of Evaluation made the following statement in its first report:

> A balance is to be achieved between, on the one hand, a precise knowledge of the objectives and analyses of government departments which will make it easier for them to identify with the conclusions reached, and, on the other, a study procedure which is not limited simply to the assumptions underlying the decision procedures.
Although the emphasis is on learning from managerial evaluations, a more detached type of evaluation makes it possible to include external viewpoints (Trosa, 1992). In the Finnish evaluation, an attempt is made to take external viewpoints into account particularly by means of surveys among members of the public, although a steering procedure closely connected with the management system is maintained. In the final conclusions phase, it will be interesting to compare the contributions made by the two approaches.

Evaluation in Finland does not provide for any scientific control mechanism. This complies with the objective of pragmatic, rapid action, that is assumed to have a strong political input. However, without this input, the evaluation conclusions might not be so valuable and instructive in the absence of a specialist methodological body -- considering that the studies most closely related to social issues are the most difficult and most delicate to handle from the methodological viewpoint. Therefore, the analysis of the views of the public refers to a comparable study which was conducted in 1987, although the conditions and approach probably differed from those of the present context.

**Follow-up to evaluation**

The follow-up is obviously the whole point of an evaluation that is mainly intended to produce rapid, wide-scale results as an input for decisions and to generate a substantial learning effect commensurate with the sponsor’s involvement. The evaluation culture in the Finnish public service is to be developed by a training cycle for public servants and the preparation of a practical guide.

The value of such an approach, which has not been sufficiently developed in France (Trosa, 1995), is that evaluation is integrated in the normal cycle of administrative activity. It would probably be advantageous if it were accompanied by an information system guaranteeing the lasting use of public service performance indicators. This aspect, which was taken into account by the evaluation sponsor in the project phase, could be further developed.

**Conclusion**

From the methodological viewpoint, the evaluation system discussed has obvious advantages which can be summarised under three main points:

- The approach closely identifies the main methodological difficulties of evaluating public management reforms, and proposes solutions marked by some pragmatism (evaluation duration and budget deliberately limited, use of existing studies some of which have already been published, acceptance of limits to scientific precision).

- The evaluation combines a fairly open field of analysis and inquiry with a strong managerial input.

- The procedure used, which involves different evaluation aims (cognitive and instrumental aspects, learning effect) will provide useful information on the relevance of the processes and objectives concerned. From this viewpoint it will make a very positive methodological contribution to future evaluations of the same type.
REFERENCES


STRATEGIC MANAGEMENT IN GOVERNMENT:
EXTENDING THE REFORM MODEL IN NEW ZEALAND

by Alex Matheson, Gerald Scanlan, and Ross Tanner
State Services Commission, New Zealand

Introduction

The idea that government can manage itself strategically, in a similar manner to well-performing private sector organisations, has featured prominently in New Zealand's contributions to PUMA work since 1994. The purpose of this paper is to outline the rationale for a strategic approach to government management, to describe the elements of the approach which we have developed, to assess its impact to date on the performance of government and to point out future opportunities and challenges.

Rationale for a Strategic Approach to Government

*Government is a complex business requiring high quality decisions.*

A critical determinant of successful government is the ability to make good decisions, and manage their implementation. Modern government is complex. It requires thousands of decisions to be taken and acted upon each day. The compound effect of these decisions is enormous, due to the size and pervasiveness of government. The dramatic decline of New Zealand's fortunes during the 1970s and early 1980s exposed the inherent flaws in our version of the traditional model of centralised government decision-making and management. We discovered the cumulative costs of poor decisions the hard way.

We learnt that, under stress, centralised authority fails to cope with entrenched features of public administration such as value conflicts, ambiguity and uncertainty over cause and effect, asymmetries of information about performance and the likely consequences of decisions, skewed incentives, and the difficulty of distinguishing at any given moment the accidental from the systemic, and the trivial from the significant.

---

10 This paper was presented by Ross Tanner, Deputy State Services Commissioner, to the OECD’s Public Management Service. It was written by Gerald Scanlan, drawing on papers prepared within the State Sector Development Branch of the State Services Commission, under the leadership of Alex Matheson.

11 In this paper, government refers to the compete system of public governance and management, including both political and administrative dimensions; Government refers solely to the Ministers who collectively exercise political authority.
The public sector reforms begun in 1986 reflected the realisation that our arrangements for government decision-making and management needed a radical overhaul. New Zealand set about developing a management system that would provide greater assurance of the quality of decision making and the overall level of performance. For high quality decisions it is important that people are appropriately empowered. It is vital that they have a good understanding about Government’s overall priorities and the types of behaviour being sought, and about the impact of their decisions and actions on the government system. It is important that good information be available to them or that they have incentives to get that information. It is also crucial that those people be held properly accountable for their decisions.

**Making managers manage improved the quality of decisions.**

The shift from centralised regulation to decentralised management was the defining event in New Zealand's public sector reform. Deregulation of the public service compelled managers to move from a passive compliance mentality to an active management mentality. The contractual emphasis in the New Zealand model pushed the need for intelligent decisions on resource allocation down through the administrative chain to the point at which services are delivered. This forced managers at every level to focus on the Government's objectives, to manage the risks to those and to become more responsive to the requirements of the recipients of their services. Freedom of information legislation, introduced prior to the reforms, was instrumental in changing attitudes to the value of information, and in improving the volume and quality of information about government. Decisions at all levels are now based on better information, better budgeting processes and a longer time frame.

The process of clarifying objectives and then demanding good quality information on their achievement illuminated what public servants actually do and assisted decisions around resource allocation, notably during a period in which government was reduced in size. However, there are limits to the gains to be had through specification and monitoring. While the reforms made government more manageable, they did not in themselves ensure that decisions affecting the longer-run interests of government as a whole were properly balanced against the particular interests of the component parts. The devolved management system tended to obscure the fact that government is a single enterprise, which should be unified around commonalties of purpose, behaviour, management standards and reputation. Re-building commitment to the collective interests of government was an early priority of the National Government after it was first elected in 1990.

**Autonomy, though necessary, is insufficient for effective governance.**

The trigger for a strategic approach to managing government was the realisation that effective public governance now requires the capacity to design and execute long-term solutions to complex problems, often with fuzzy boundaries and both international and inter-generational consequences.

It has become a truism of business that the increasingly integrated global economy is testing the competence of managers to the limit. Government, too, operates in an international market for ideas, people, reputation, information and policies. Democratisation, the information revolution and global

---

12 New Zealand held its first general election under the new proportional electoral system (MMP) on 12 October 1996. At the time this paper was presented to the PUMA meeting in early November, the identity of the new Government had not been determined. On 11 December, a coalition Government was formed in which the National party is the majority partner.
integration have reduced the capacity of governments to control conditions within their borders relative to the international community, while increasing the relative autonomy of their citizens and raising their expectations of government.

The only decision-making model which can cope with these conditions is one based on directional signals about the Government's purpose and priorities, which guide decisions taken autonomously by managers close to the action. Good decisions then depend on a high-quality flow of information, well-tested analysis and advice, and a balanced set of controls, plus the ability to learn and adapt. We can capture these insights about the modern conditions for optimal performance in the motif: direction, autonomy and control. Implicit in this formulation is the idea of dynamic equilibrium between the three elements.

Making governance more effective challenges the role of government.

In government as in business, one of the greatest benefits arising from the adoption of a strategic approach to decision-making is the premium it places on identifying and strengthening core competence. It leads to a disturbing question: what is government good for? In the initial period of reform, the question was posed in the negative: what is government not good for, and therefore should no longer do? Now we are better placed to identify the important and unique job that government has to do, and has to be good at doing. Aberbach and Rockman\(^{13}\) propose that government's essential competence lies in its ability to:

- make quality long-term decisions;
- create and distribute knowledge;
- implement decisions effectively; and
- mediate amongst competing interests.

Strategic government also means good government, since the quality of government will increasingly determine international competitive advantage. In the words of one commentator looking to the requirements for national success in the next century:

> While efficiency is a necessary condition of wider prosperity and influence, it is not sufficient. Enduring prosperity requires societies that are safe, ordered and honest.\(^{14}\)

The importance of leadership

The strong need for developing a strategic approach to government in New Zealand was matched by the readiness of Ministers and senior public servants to take the reform model into a new phase. Political leadership was vital. The National Government made a public commitment to a range of policy objectives over the period up to the year 2010. It also decided to publish annual updates of progress towards those objectives. Some Ministers adopted the same approach within their portfolios, and published '2010' documents for areas such as the environment, science and technology and education. It is difficult to overstate the importance of such clear and public objectives, and of sustained commitment to their achievement.

---


Dr Allen Schick, Professor of Public Policy at the University of Maryland and Visiting Fellow at the Brookings Institution, recently concluded a commissioned, independent study of the New Zealand public management reforms. He found that few elements of the New Zealand reforms have been as well received by Ministers and senior public servants as the recent initiatives to strengthen our strategic capacity. What those initiatives are is the subject of the next section.

The Elements of the New Zealand Approach to Strategic Management

**Designing for selectivity and simplicity**

Being strategic means being selective -- sorting the critical few from the important many, and giving that selection 'bite' by shifting resources and demanding performance sufficient to make the desired difference. New Zealand opted for a working model of strategic management, whose key design idea was simplicity. Government officials wanted to minimise impositions and let improved information carry most of the weight, in recognition of the critical role of information in guiding behaviour in a free-to-manage situation. The key elements of the design are:

- a selective set of generalised, cross-portfolio policy objectives set by Cabinet (**SRAs - strategic result areas**);
- a process for co-ordinating departmental contributions to those objectives and making related resourcing decisions (**strategic dialogue**);
- a set of critical medium-term commitments (**KRAs - key result areas**), which anchor departments' strategic contributions to the policy objectives, through incorporation in the chief executive performance agreement;
- a requirement that chief executives regularly report progress on those commitments to their Minister and to the State Services Commission; and
- an expectation that chief executives will take responsibility for making, and taking care of, the connections between their commitments and those of other chief executives, while also ensuring that their own commitments flow down through their department's management chain.

The distinguishing features of SRAs and KRAs are their selectivity and flexibility. For all of the reasons that doom long-term central planning, particularly the complexity and the uncertainty of the future, it is not possible to come up with a complete set of higher level strategic objectives, finely specified. The best we can do is to pick out a few objectives which, given our understanding of the systems in which they are embedded, are most likely to exert the desired leverage on those systems. This means that most higher level output objectives are not included. Moreover, the path to achieving the selected objectives is not pre-determined. SRAs and KRAs are not so much performance checkpoints as beacons: luminous catalysts of information which guide behaviour and focus attention in a devolved management situation.

Creating common space between politicians and administrators

We have coined the phrase "the purple zone" to describe how the approach works (refer diagram at Annex One). The traditional bureaucratic model could be described as an 'indigo' zone: one in which it was impossible to distinguish the actions and responsibilities of politicians and administrators, and in which information about cause and effect, expenditure and result was elusive and vague.

The initial period of reform resulted in a dramatic separation of primary colours: blue for politics, and red for administration. Between them rose a wall of contracts, in which the resources were exchanged for results. But experience taught us that much of what is important about government is not amenable to close specification and monitoring, and does not fit neatly into annual cycles and is not easily separated into compartments labelled politics and administration. The purple zone is the product of that experience: an amalgam of separation and integration, in which the benefits arising from distinctive responsibilities are complemented by those flowing from strategic coherence and common "branding".

Essentially, the purple zone is the arena of conversation between Ministers and their senior officials which ensures that appropriate attention is being paid to the Government's priorities and that the consequent interdependencies, risks and departmental initiatives are being managed in a way that serves the interests of government as a whole. Its tangible products are the SRAs and the KRAs around which the public service builds its purchase and performance agreements. The conversation also strengthens the informal, cultural components of strategy, notably the sense of shared purpose and collegiality among the key players. The red and blue zones, and the purple zone in which they blend, encapsulate the total performance management system. The trick from an overall system perspective is to maintain an optimal balance between direction, control and the autonomous energy of individual managers.

An incremental design

It is important to understand the evolutionary nature of our approach. It was not the product of a master plan devised by Ministers or officials. Rather, the core elements emerged separately and came together in a loose, experimental manner.

The initial step was taken by the Government, which in 1993 published a long-term vision for New Zealand through to 2010. The strategic management process gained momentum the following year as a result of the Prime Minister's concern that the public service give more tangible expression to the medium term steps it would take to move toward the desired vision. We developed Strategic Result Areas (SRAs) - a limited number of major desired results for the public service as a whole. The SRAs inform the agreement of Key Result Areas (KRAs) for chief executives of individual departments. The critical difference between the SRAs and the KRAs is that the latter are medium term, largely within the influence of the public service and formed within the assumptions of the budget policy statement. Resources are recognised to be limited and trade-offs are made. Annex Two contains examples of current SRAs and KRAs.

KRAs provide the means by which departments expose the 'critical few' targets which they wish to achieve over a two to three year period. Measurable milestones associated with each KRA allow Ministers and others to track progress towards those targets. KRAs might have a direct or indirect link to an SRA, or they might focus on an important aspect of organisational capability. KRAs should be few in number - generally not more than six. The requirements to expose KRAs to critical scrutiny by central agencies and other departments with related interests, and to obtain the Minister's endorsement of the selection, ensure that chief executives direct effort and resources towards those few priorities which offer
the greatest potential return to the Government. Moreover, chief executives have a direct stake in the strategic performance of their department through the incorporation of the KRAs in their performance agreement with the Minister.

*Conversation in the common space*

The annual SRA/KRA process encourages ministers to interact with senior public servants on the desired high-level deliverables for the public service, and for departments and central agencies to interact with each other to address cross-sectoral issues, establish priorities and assist Ministers to fix key responsibilities for individual department delivery. The process sets individualised service delivery within a long term and whole-of-government context.

The key instrument for Government strategy is, and has always been, the budget process. The new strategic approach seeks to enrich the budget process by ensuring that wider policy issues are more consciously reflected in the formation of the budget and in appropriation decisions. The present Government has instituted an annual retreat to allow Ministers to more deliberately set the budget parameters and their strategic objectives within them.

*Distinguishing substance from shadow:*

Because the SRA/KRA framework differs in design from the established contractual instruments (the purchase and performance agreements), we have found it necessary to explain what the strategic approach is *not* about.

*...not strategy by measurement...*

The approach is not is an elevated form of management by objectives; it is not a planning model, but rather a strategic decision support model. Those who view it as a planning tool tend to advocate ever tighter pre-specification of intended performance. Effectively this means trying to treat strategic objectives in the same manner that annual operational objectives are treated in the purchase agreement. Should this view prevail, and it is one which comes very naturally to public bureaucracies, we risk undermining the advantages we have gained from letting managers manage, and imposing progressively higher transaction costs in the form of more detailed planning, monitoring and reporting documentation. We would also shift the risk of not doing the right things from departments to the central agencies, and encourage risk-averse target setting.

Key Results are (or should be) inherently quite difficult to achieve. Their accomplishment will often be subject to unpredictable influences. This means that total pre-specification will rarely be possible, much less desirable. Much must be left to departmental discretion, while the evaluation of performance against KRAs tends to contain a considerable element of 'judgement', related to the particular circumstances and track record of each department. While there is much that can be done to improve selection of the few key things which will most affect department performance, the notion that *generically* KRAs can be improved by more pre-specification is flawed.
...nor strategy by inclusion.

A second possible distortion is the "me too" syndrome. Those whose function is not captured by an SRA or a KRA feel that the function is somehow being treated as of less worth, and will push for its inclusion. While the argument has a superficial appeal, conceding it would destroy the strategic character of the model. Given the role of SRAs as a selective set of major cross-sectoral results, it is to be expected that many departments will not relate directly to SRAs, and many functions will not be captured in their KRAs. Indeed, the more comprehensive SRAs and KRAs become, the less valuable they are as a strategic management tool, given the limits to the kind of influence leaders of organisations can bring to bear at any one time.

A related distortion is that of special pleading. It is tempting for advocates of particular causes to seek the explicit inclusion of their cause in the KRAs, or even better in the SRAs. This distortion stems from a failure to distinguish between the wide category of things which are important to government, and the selected high level results judged most likely to further the Government's overall strategy over a specified period. It is possible to say that while $x$ is a very important component of government policy, $x$ should not be an SRA. Otherwise, special interests will rapidly degrade SRAs into a shopping list of unsorted priorities.

The final selection of SRAs is not a measure of importance per se but a dynamic judgement about what key points of leverage should receive scarce high level attention, and equally scarce additional resources, over a defined planning period. Because something is an SRA for one period does not mean it should continue to be so. Indeed, if our strategising works we would expect a healthy level of turnover, as attention moves from one major problem to another.

Interaction of formal and informal processes

The new strategic management system is straightforward -- Ministers decide and specify the Government's priorities, the public service distils these into achievable objectives for each chief executive and department, Ministers and chief executives conclude formal agreements to cover these, performance against the agreements is assessed, and the information obtained feeds back into the loop to improve the quality of the next cycle.

The rather simple formal strategic mechanism is wrapped in an informal process of conversation, monitoring, adjustment and evaluation. The informal process cannot readily be described or specified, but it is critical to making the approach work.

It is important to restate that the process adopted is not geared towards producing a centralised strategic plan which is circulated through the public service, nor creating a set of iron-clad undertakings which consume a chief executive's every waking moment. Rather, it is geared to constructing an 'open agenda' between Ministers, chief executives and the central agencies which ensures that major issues receive regular, detailed and well informed attention and discussion. Clarity about what sits in the common space between the key decision makers then should lead to a common view of the nation's medium-term direction, as a basis for good decisions and performance.

The next section of the paper examines the impact of the new system on government's performance so far, and on the roles and perspectives of the key players.
Impact on Performance

We introduced a strategic management process to improve attention to longer-term priorities, government co-ordination and consideration of overall effectiveness, which were initially overlooked in the shift away from central regulation. But what results can we claim so far from our attempts to re-balance our devolved management system?

Departments are becoming more strategic.

There has been a considerable spin-off from the Government's strategic approach within departments. In particular, departments can now make their Budget bids and do their business planning within a common strategic framework. Almost all departments are developing strategic business plans, allowing them to make clear connections between what they have to do to contribute to the wider strategic picture and the organisational implications in terms of structure, skills, capital investment and external relationships. They are also more able to involve managers and staff in their own strategic management initiatives, such as flattening structures, devolving managerial responsibility, empowering front-line staff, investing in information technology, and paying more attention to evaluating the impacts of policy and services.

Departments are beginning to co-ordinate business strategy.

Perhaps the most promising development is the emergence of sectoral strategy. Given the way in which strategic result areas are framed: high-level, medium term and cross-portfolio, sectoral co-ordination is essential if strategic objectives are to be achieved.

Because it is largely unmapped territory, the emergence of sectoral strategy has prompted some quite innovative approaches. The environmental agencies have formulated a "green package" of priorities and advocated an "environmental envelope" of new spending to achieve them. Research and technology related agencies are working towards a GDP-linked spending target, and are trying to ensure that all relevant expenditure in department budgets is identified for the purposes of counting against that target. In the area of border control and biosecurity, effort has been applied to developing sectoral policy principles which can be consistently applied through the key result areas for individual departments.

The development of sectoral strategy has already led to improved information flows, more substantial consultation with commercial interests and non-profit organisations, greater clarity about the Government's vision and priorities and a surprisingly powerful synergy among agencies with a history of sometimes fractious relationships.

Chief executives now have a bigger job.

Chief executives have become the Government's business partners, responsible both for helping to shape the broader policy environment and for making real progress towards specific objectives. Their role compels them to contribute to the spectrum of the Government's interests, beyond their organisational boundaries. The specification for the ideal chief executive is changing, favouring persons who can see the big picture and convey that to their organisation, who have an entrepreneurial bent, who feel comfortable in common space and are able to make connections between their business needs and that space. Departmental managers are also becoming more demanding of the government system as a whole, now
that their sights have been raised. They expect a more sophisticated approach to government strategising, budget setting, policy resolution, interactions with central agencies and Parliamentary accountability.

A reoriented role for the central agencies

The story of the New Zealand reforms can be expressed as a swing from central regulation (control on inputs) to self-regulation (control on outputs) and now to system regulation (control on strategic capacity and results). Pivotal to this story is the role of our three central agencies, the Treasury, the State Services Commission and the Department of Prime Minister and Cabinet, who have both separate and shared responsibilities for the government-wide functions of financial management, policy co-ordination and performance assurance. The three agencies form the Government’s ‘corporate office’, a term which conveys a sense of the fuzzy boundaries and over-lapping responsibilities which the three roles entail. Another way of viewing the three roles is that they span the Government's 'value chain': resource inputs, production process and strategic results.

Trends in the central agency role include a more consistent focus on the whole business of government, pooling of expertise, a greater concern for adding value through knowledge creation and keeping up with the play internationally, and a preference for facilitation and guidance over direct intervention. In particular, the three central agencies are seeking improved alignment amongst the instruments for which they are responsible, namely the Budget process, Cabinet decision-making and chief executive appointments, contracts and performance assessment. Integration of those instruments, underpinned by shared values and confident transactions amongst the agencies, should help departments manage in the wider public interest without undue central control.

Strategic considerations underpin the Government's ownership interest.

We have also put considerable effort in recent years into clarifying the interests of the Government as owner of its departments, as distinct from the earlier focus on its interests as the purchaser of their services. The work on clarifying the ownership interest, and developing an approach to managing that consistent with our devolved management system, has neatly complemented the greater accent on strategy. We have determined that the owner is interested in strategic alignment, commitment to collective interests, future capability and long-run cost effectiveness, with the following implications:

- The Government is a single owner, with a common strategy and co-ordinated arrangements for making important decisions.
- The Government has a strong interest in ensuring it has the right set of institutions to serve its priorities, and in ensuring they are capable of meeting future demands, reliably and cost-effectively.
- The Government can impose collective obligations and standards of management and behaviour on departments and other Crown organisations, so that the actions of one do not have adverse consequences for others, or create unnecessary costs between departments, or damage the reputation of government as a whole.
- Managing the Government's ownership interest in a way that gives Ministers confidence that their department is in good shape and is making the desired impact on the relevant priorities is the cornerstone of the chief executive's job.
For chief executives to meet Ministers' requirements, they need to: operate reliable systems for managing resources and activities; place effective controls on those systems; select and develop people with the right skills; develop a capable senior management team; and ensure that the right information and incentives are in place to align the efforts of all parts of the organisation with its purpose and the relevant priorities of the Government. They also need to co-ordinate their business strategy with that of other departments with overlapping contributions to one or more of the Government's priorities.

**The strategic approach suits both ownership and purchase interests.**

There is considerable overlap between the approaches taken to our strategic purchase process and to managing the ownership interest. The first, and probably most important, common feature involves creating a context for conversation between Ministers and chief executives about matters of importance to the owner. There is now an expectation that chief executives will be able to offer Ministers reliable information about their management of the Government's interests as owner, and to identify the ownership implications of Ministers' major purchasing preferences.

The second common feature is the provision for Ministers to require chief executives to develop one or more key result areas (KRAs) which address aspects of the ownership interest about which Ministers have concerns. Chief executives then have to provide credible reports on their progress.

The third common feature relates to the strategic co-ordination function of the central agencies. At key decision points in the life of a department, there will be an opportunity for Ministers and central agencies to test the quality of its management of the owner's interest, and to subject its business strategy to critical scrutiny. Such points include bids for new capital or for a change in output price, or the transition to a new Government, new Minister or new chief executive.

**The new role for control: balancing autonomy and direction**

Adopting a strategic process has also altered our perspective on control, consistent with the need to balance that with direction and managerial autonomy. Much of our new thinking on control has been prompted by Robert Simons' work on a balanced approach to control. Simons\(^{16}\) has identified four complementary levers or modes of control:

- **Reflective/integrative** control, through attention to common interests, values and beliefs. This is the prime means of influence provided by leadership.

- **Pre-emptive** control, through specification and enforcement of boundaries, imperatives and prohibitions. Successful organisations tend to make explicit what behaviour is unacceptable, and seek to reinforce this in the minds of staff by swift and certain action against transgressors.

- **Diagnostic** control, through clarification of targets and monitoring of their achievement. Most financial, output and process management information systems support diagnostic control.

---

Interactive control, through open-ended dialogue geared to adjusting strategy as events unfold future. Inherent in this is the notion of strategy not as a single plan but as a stream of adaptive decisions.

The old public service was run predominantly by integrative controls (culture) and pre-emptive controls (the centralised regulation of inputs and process). Great gains were achieved when detailed pre-emptive controls largely gave way to diagnostic controls: making managers manage and introducing output based appropriation and monitoring to reduce cost and improve efficiency. However, once government’s attention moved towards longer term strategy, this set of techniques was clearly insufficient. Our search for a point of balance between the available means of influencing the performance of government has now led to the introduction of an interactive control system -- the strategic management process.

The advantage of a multi-layered approach to control in a devolved management system is that it allows the mix of controls to be tailored to differing circumstances and risks. The more that can be achieved through informal means, such as affirmation of values and beliefs, development of relationships and opportunities for dialogue, the less reliance need be placed on formal instruments and exchanges. By shifting the balance, from formal to informal and from extrinsic to intrinsic, the control system can send a powerful signal about the role of trust in a well-performing management system.

Success in staking common ground

Our reflection then is that there have been considerable gains from the original decision to adopt a more strategic approach to the business of government. The success to date of the New Zealand approach can be attributed to the willingness of the major players -- Ministers, chief executives and central agencies -- to take shared responsibility for making it work. Reinforcing that responsibility is a range of influences, both subtle and sharp, which keep the public service focused on strategy. The sharp influence of the pressure on chief executives to perform means that the Government's priorities receive attention and ensures that the chief executive's commitments are meaningful. More subtle influences are provided by the shared values and web of relationships found in the public service, which foster a culture of cooperation and sustain the 'free play' needed to resolve conflict and adjust quickly to changing circumstances.

But while the successes have been real and substantial, challenges and opportunities abound which will test our willingness and ability to sustain strategic momentum. The final section of the paper discusses the most prominent of those challenges and opportunities.

Challenges and Opportunities

Improving the process

It is important to acknowledge that our strategic management process, based on the SRA/KRA framework, is probably somewhere between infancy and adolescence. Our first set of SRAs are not uniformly well-sculpted objectives. KRAs, too, are of uneven quality. Improving them will require more careful analysis and challenge from central agencies and departments around what are the key operational problems for the department involved, better tailoring of specification effort to operational risk and more attention to the quality of reporting and its analysis.
However, this is an area where the **process** of focusing on the major issues, and seeking alignment amongst the key players, is more important than the final written **product**. Accordingly, care must be taken that this process, designed to facilitate debate and the making of joint cause between senior decision-makers, does not regress into centralised regulation with its associated bureaucracy, compliance costs and constrained operational autonomy. We can progressively improve the process by ensuring that it is supported by good quality staff, that it utilises good quality empirical information and evaluation relating to progress against the Government's priorities, and that the conversation is not unduly impeded by institutional rivalries or historical baggage. The trick is to maintain a balance among the various means of influencing strategy in a devolved management system (Simons' four levers above), suited to the prevailing circumstances.

So far, the strategic management process has won good support from the key players, as Professor Schick has confirmed. There is no dispute that strategic management is an appropriate tool of government. There are evident benefits in terms of direction, focus and cohesion within government, and better information about the intentions and impact of the Government for its public and its international audience.

However, the strategic management model confronts us quite starkly with the challenge of what Aberbach and Rockman and others have identified as the things that government must do well to be effective.

**Extending governance beyond government**

The first thing that Government must do is exercise leadership, but this no longer means having a finger in every pie and solving every problem. Under the new model of governance being developed in New Zealand and elsewhere, the Government's leadership role is about framing a vision, setting broad directions, and then creating the conditions in which society as a whole can steer in the desired direction. To maintain legitimacy, democratic government must allow the wider society to share in leadership. The constitutional separation of power, pluralistic politics, a flourishing civic society, a free and competent media, demanding and informed customers for government services, academic freedom and an open economy all help keep government under control.

**Restraint: making space for other sources of leadership**

Government under control is one which exercises self-control. A prime example is provided by the Fiscal Responsibility Act. This Act takes a principle with which it is hard to disagree, namely responsible management of public funds, and spins it out into a simple but forceful form of self denial. The Act itself sets no numerical targets and fixes no immutable boundaries, but compels the Government to face the music should it step outside the zone of acceptable fiscal conduct which exists in the minds of informed commentators, whether locally or overseas. Therein lies the elegance of the Act: the Government is not the arbiter of its own fiscal responsibility; the Government no longer exclusively “owns” fiscal policy.

At a more modest level, the same principle of restraint underpins the government management system. For a variety of reasons -- constitutional and motivational in particular -- it is vital that Ministers, the State Services Commissioner and central agencies respect the managerial autonomy of chief executives. Chief executives can only learn to be responsible by having to exercise responsibility; they can only share in leadership by exercising leadership.
Improving government: becoming a learning organisation

Improving the quality of government's role depends on enriching the flow of information and analysis upon which government can base good decisions, implement them effectively and intervene appropriately to reconcile competing interests. Essentially this is a question of government's knowledge capacity: how well it learns and adapts in response to information about the impact of past or current decisions and actions.

This is a particular challenge for New Zealand, because knowledge capacity takes us into the realm of evaluating outcomes, which were quite consciously sidelined in the original reform model which assigned responsibility for outcomes to Ministers. Yet we now recognise that good strategy, and intelligent government, is impossible without a steady stream of empirical findings on whether the Government is attaining its goals, leading to corrections to strategy or management controls.

SRAs are of value in identifying, and lining the public service up behind, some over-arching operational priorities. They contribute to goal congruence, and they reach across from the political to the operational domain by influencing the goals and incentives of chief executives. However the SRAs are still expressed as high level operational goals, so far unsupported by any requirement to identify indicators of effectiveness or to have that effectiveness evaluated.

Remedying the deficiency of outcome information in our system will require a marked shift in our thinking. We need to put aside the preoccupation with accountability which has blocked attention to outcomes and agree that while nobody can be held directly responsible for outcomes, the success of the governmental process depends on each decision maker in the chain being as well informed as possible about the impact of their decisions on the community. While the output/outcome distinction remains important in terms of clarifying what public servants actually do, and should be held accountable for, externalisation of outcomes denies the reality of managing difficult functions in a complex environment and prevents government acquiring the habits of a learning organisation.

Interestingly, Professor Schick's report argues that we should cease talking of outcomes as things which politicians can be responsible for. He suggests that outcomes...

...should be employed more for formulating policy than for maintaining accountability.
They are powerful directional signals: are things getting better or worse?

The approach to evaluation that best fits a system of both democratic governance and devolved government management is one where important empirical and theoretical information is produced independently and made available as a matter of course to policy makers, producers and consumers. The Reserve Bank and Fiscal Responsibility Acts have demonstrated the benefits of making good quality information widely and routinely available as a means of modifying behaviour and encouraging appropriate decisions. There is no reason why we cannot gain similar benefits from a commitment to creating good information about other aspects of government activity, whether through independent evaluators or the mandatory production and disclosure of key outcome information. This will help 'close the loop' on government's performance, aid judgements about benefits gained for money spent and strengthen our strategic capacity.

The gradual strengthening of strategic management in recent years has helped create the demand conditions for outcome information that were previously missing. Now that Ministers are more strategically focused, they are seeking better advice and information on the impacts of policies than they currently receive.
A complementary managerial strategy would give attention to the skills, incentives and attitudes needed to ensure officials seek out and respond appropriately to evaluative outcome information, including the occasional 'forensic' studies which extract lessons from instances of policy or operational failure. Various quality-based initiatives such as awards, league tables and charters have been used elsewhere to encourage a culture of continuous improvement in government. It might be timely for New Zealand to look more closely at that type of incentive.

...and strengthening our institutions

New Zealand's public sector reform programme has resulted in a large number of small departments, each of which must sustain a fairly standard load of management and accountability requirements. Concerns have emerged that in some departments, controls are draining managers' energy and narrowing their perspective. We are now focusing attention on ways of reducing systemic obstacles to dealing with the complex problems faced by government. In particular, we need to strengthen our capacity for co-ordinating departmental policy initiatives and dealing with unintended effects. We also need to improve our understanding of how the operation of the public sector labour market affects institutional capability, and what interventions might be appropriate to offset adverse features of that market.

Stronger institutions are needed to sustain a shift in the Government's and community's priorities from macro-economic to macro-social policy. That shift will test the Government's willingness to allow community interests to take the policy initiative and to support devolved decision-making with resources and information. We need to get closer to community interests to build our knowledge about the constraints on, and effects of, social policy choices.

The shift will also test public confidence in government, based on the integrity of government institutions and the quality of their services. Public service managers have to temper their operational autonomy with a strong sense of ethics and professionalism, while meeting the steadily increasing service expectations of the public.

The challenge of political change

The final challenge is that posed by our transition to proportional representation and the likelihood of consequent changes in government processes. Under the MMP electoral system, public servants will have to be more skilled in the selection and management of processes of decision making. Policy processes are likely to be more open and require more public input, multi-party consultation and Parliamentary debate. The indications from elsewhere are that the new political system will place a premium on clear Government objectives, and will create a demand for good information about whether the Government is realising those objectives.

Conclusion

According to Professor Schick, the recent advances made in building New Zealand's strategic capacity "demonstrate the elasticity of the reforms and their openness to new ideas,[since the advances have been] accomplished in ways that comport with the logic and practice of the New Zealand model". We have experienced innovation because we have permitted innovation.

While the strategic management processes discussed in this paper represent a logical extension of the reforms, they are now taking us into unfamiliar territory. We have created an arena in which strategy
can be debated, priorities tested, outcomes evaluated and significant benefit/cost assessments reached. But we are not yet sure who all the players will be or what rules they will play to. What happens will depend on the quality and motivation of the players, their capacity to communicate with each other and the depth of their shared knowledge. There is no more important determinant of the success of New Zealand’s public policy than the quality of this game; played well, it will be one from which we all emerge as winners.
ANNEX ONE
Strategic Management Model
The Purple Zone

Integration of Budget and Wider Strategy Formation Process

Strengthened “Strategic Conversation”

Interactive Political/Administrative Decision-Making

Outcome Information needed to fuel Strategic Decision Making
ANNEX TWO

Strategic Result Areas for the Public Sector 1994 - 1997

1 maintaining and accelerating economic growth  
2 enterprise and innovation  
3 external linkages  
4 education and training  
5 community security  
6 social assistance  
7 health and disability services  
8 Treaty claims settlement  
9 protecting and enhancing the environment

The details of the 'Education and Training' Strategic Result Area are shown in the box below. The SRA has significant implications for several departments, including Education, Labour, Youth Affairs, Women's Affairs and Maori Development, as well as the three central agencies. All of these departments contribute to SRAs as well.

4 Education and Training

*Progress towards higher and more appropriate skill development to support the achievement of stronger employment and income growth.*

Particular emphasis will be placed on:

Development of programmes and a curriculum that will enable an increasing proportion of children to receive effective early childhood care and education, particularly those at risk.

Completion of a redesigned compulsory curriculum which focuses on generic skills, sets learning objectives and monitors the performance of students; monitoring the performance of providers and provider institutions.

Further implementation of the National Qualifications Framework to enable students to build recognised qualifications; and integration of all post-core education and training in a seamless continuum of learning opportunity.

Ensuring that more New Zealanders have access to education and training, particularly through Skill New Zealand; development and implementation of systems that will enable full participation in workplace and post school education and training which leads to recognised qualifications.

Improved systems for the effective delivery of resources to students and providers to enable the flexible provision of educational services to the community and, in particular, to target groups within it.
Key Result Areas

The KRAs agreed this year between the Minister of State Services and the State Services Commissioner, as chief executive of the State Services Commission, include:

The Commissioner will ensure the Public Service completes an effective transition to the new political environment by 30 June 1997.

The Commissioner will put into place, by 30 June 1997, a chief executive employment framework in accordance with the Government's strategic objectives for the Public Service.

The Commissioner will identify the Government's key ownership interests and risks in the management of the Public Service.

The Commissioner will provide assurance and advice to the Government on performance relating to its key ownership interests across the Public Service by 30 June 1997.

The Commissioner will work with other relevant chief executives to implement the Government's objective of strengthening the management and accountability of school boards of trustees and tertiary councils.

The Commissioner will ensure that the on-going role of the Commissioner beyond 1996 is fully defined and that the Commission has the capacity to support this role.
SOURCE DOCUMENTS PRODUCED WITHIN THE STATE SERVICES COMMISSION


MATHESON, A., Strategic Management of Salamanders and SRAs, June 1996.

MATHESON, A., Programme Evaluation: Something Funny Happened to Outcomes on the way to the Reforms, August 1996.


SCANLAN, G., Public Service Management Framework (Draft), May 1996.
In this commentary I will discuss: (1) the evolutionary nature of performance management systems and reform efforts, (2) the importance of a contextualistic perspective in reforms, and (3) the key components of New Zealand’s strategic management approach.

The Evolutionary Nature of Performance Management Systems and Reform Efforts

Performance management, in a broad context, can be viewed as a selective learning model. More specifically, as a “Darwinistic model” whereby leaders and others select best practices or approaches and change or dismiss ineffective ones. The term Darwinistic is used here as an analogy from biology; however, in a performance management framework the selection process is, or can be, more studied and based upon conscious and selective choices. This is because the most basic concept of performance management is that of leadership. The leaders must have a clear sense of direction, some good measures or navigation tools, and a systematic forum, for applying various change management tools, to (a) select and strengthen best practices or approaches toward a particular direction and (b) change or dismiss ineffective practices.

However, leaders frequently need to reassess their direction in the context of unforeseen external and internal changes. Managing change requires the use of flexible and selective management tools. In line with these principles for managing change, the paper on New Zealand experience offers a description of strategic management as a flexible, selective and contextual management practice.

The Importance of a Contextualistic Perspective

It is important for leaders of future reform efforts to learn from others. However, while learning from the New Zealand experience, its historic and current context needs to be taken into consideration. Contextualistic world views were first explained by S. C. Pepper in 1942. Others have furthered Peppers work to help us appreciate the complexities associated with understanding the critical components of social changes. Contextualism has mainly two lessons that may apply here.
One lesson is that reforms do not happen in vacuum and a holistic approach to the efforts must be applied for co-ordination. The individual components of reform efforts must be synchronised toward a common future direction. Reforms have involved many different strategies, for example, privatisation, decentralisation, increased accountability and flexibility, empowerment of employees, quality management efforts and more. These change strategies need to be synchronised at the highest level to avoid creating inconsistencies.

Another lesson from contextualism deals with the different historic context and current context of reforms. The New Zealand reforms had a certain historic context that was unique. When individual components of reforms, in any country, are studied, surprising variables and combination of variables may account for much the successes and failures of the efforts. For example, variables such as the power of the individual players and the interaction among them, the willingness of a politician to focus long-term in the face of certain short-term contingencies in the political environment, special interests, the media, unions, and more. The appreciation of historic and current contexts, gives us, the students of New Zealand’s reform, a dose of reality while questioning how best to apply the lessons learned in different environments.

New Zealand’s Strategic Management Approach

The key components of strategic management include: (a) long-term focus, (b) co-ordination of departmental contribution to long-term priorities, (c) a common space between politicians and administrators, and (d) the focus on effectiveness and therefore commitment to managing outcomes. The New Zealand paper acknowledges that these components were overlooked in New Zealand’s initial reform efforts, but its extended reform programme has sought to develop these principles within the management framework. Indeed, the New Zealand paper describes its strategic management process as a contextual, flexible and selective change management approach. Its strategy for integrating each of these components is examined in turn.

A long-term focus was developed through the use of strategic results areas (SRAs) and key results areas (KRAs). The New Zealand Cabinet uses the SRAs to identify a selective set of generalised, cross-portfolio policy objectives. This results in a strategic long-term focus on a few critically important policy areas. Medium term commitments in the form of KRAs are then set within the context of the SRAs. The KRAs are set by departments and tie in with the performance agreement of the chief executive.

New Zealand’s long-term focus for the country was expressed in Vision 2010, or Path 2010. Subsequently, the SRAs provided the forum for prioritising efforts for future accomplishments of government departments and agencies. However, it is of interest to see what will happen to Vision 2010 and the set of SRAs when new leaders take over, leaders who may not have participated in formulating Vision 2010 and setting SRA priorities. Long-term targets may need to be set cross-politically to avoid the risk of a new government ignoring the long-term goals set by the previous one. This may have significant implications for the future evolution of political systems.

The co-ordination of departmental contributions to long-term priorities is now being achieved through the SRAs. SRAs are developed through a strategic dialogue with the understanding that most SRAs are cross-departmental. This makes sense, since real world issues are departmentalised. For example, efforts to reduce crime rates may have implication for many departments, such as education, labour, health and justice.
The leaders of the Total Quality Management (TQM) paradigm (e.g. W. Edwards Deming and Jo Juran), centred their philosophy on consistency of purpose in achieving long-term results and linking together the often disparate procedures and processes that develop within and between organisational units. Their approach was contextualistic and centred on leadership and continuous improvement in processes, services and products. Many people doubt that our current political environment, often with elections every few years, will allow the principles of TQM to be effectively applied in the public sector. The verdict has not been reached, yet the nature of New Zealand’s strategic management with its cross-departmental co-operation, and built in flexibility, may serve as a good basis for quality management efforts.

The New Zealand paper develops the concept of a “purple zone” to describe the common space between politicians and administrators. It is in the purple zone where the SRAs are systematically broken down to KRAs with input from both politicians and administrators. In understanding management strategies, the notion of a purple zone is appealing in its potential as a forum for more constructive and trusting relationships between politicians and administrators. However, it remains to be seen whether this forum helps overcome the temptation among politicians to provide detailed political supervision, or interference, once targets have been set. The purple zone in New Zealand provides an opportunity for preventive action and learning for both politicians and administrators. It will be interesting to see how this forum will evolve in the future.

Strategic management provides a focus on effectiveness and outcomes. Where it is effective, it should help leaders manage for results. New Zealand’s reform efforts have centred on outputs, rather than outcomes, as the means for assigning accountability and assessing results.

The challenge associated with managing for outcomes is that outcomes often take a long time to materialise. For example, if policy-makers set a long-term outcome goal to reduce crime rates, it may take years for outcome measures to reflect improvements. Thus, the more immediate focus on the outputs produced in government has been an easier way to deal with accountabilities and performance agreements. Today, policy-makers are moving closer to accepting that outcome goals and measures are needed in order to manage for results. A way around the output-outcome dilemma is to identify, select and manage the outputs that most closely correlated with the desired outcomes. From a contextualistic perspective, such a process would start with setting long-term targets, clarifying outcome goals and subsequently breaking them into output and process goals. Thus, going from a holistic perspective (i.e. a big picture perspective) all the way down to the unit level or even to the level of individual performers. Successful use of outputs as an interim management focus (as in New Zealand) relies on the ability to identify and select the outputs that are highly correlated with producing the desired outcomes.

Conclusions

The strategic management approach described in the New Zealand paper is contextual, flexible and selective. It appears to be a good steering system for continuous improvement and management for results in government. New Zealand’s experience provides many countries with an opportunity to learn best practices in the area of strategic management, and it also furthers speculation about future trends in public sector management reforms.
THE CONTRACT MANAGEMENT PROJECT IN DENMARK

by Peter Kjaersgaard Pedersen,
Hanne Dorthe Sørensen and Jakob Buhl Vestergaard,
Ministry of Finance, Denmark

Introduction

This paper discusses the Danish experience with contract management. Contract management has been in operation since 1992 and has been an important instrument for the implementation of performance management in the Danish state. The basic principle of the system lies in the establishment of a contract between a department and an agency defining targets and required results. The paper describes the position of contract management in Danish administrative policy tradition, the background to project implementation, experiences with contract management, and the problems and possibilities associated with the system.

Project Background

The need for increased focus on results

The contract management project has its background in the development of expenditure control policy in Denmark. In common with the other OECD Member countries, Denmark experienced explosive public-sector growth during the 1960s and 1970s. This led at an early juncture to deliberations on the ways in which a control of public expenditure might most effectively be implemented. The method adopted in Denmark consisted mainly of ongoing reforms to the budgetary system based on the concept of "framework budgeting". Under this concept, the budgetary system is structured as a set of top-down budgetary frameworks established on all levels (starting with the ministry as a whole on down to the individual agency). This improves policy prioritisation between principal objectives and ensures tight control on expenditure policy. In the wake of the introduction of framework budgeting, both ministries and agencies were given increased freedom of action, so that over the years a substantial decentralisation of responsibility and competence has taken place.

The weakness of framework budgeting is that it is a solely input-oriented system, not connected to the services that agencies provide. Therefore, in the extreme situation it is possible to imagine an agency that meets the requirements of its budgetary framework, but that is both low on productivity and that does not provide services perceived as relevant by its users. The weaknesses of the framework management concept have meant that the introduction of a supplementary approach providing increased focus on results has been a high-priority task for Denmark's Ministry of Finance in recent years. This supplementary approach is contract management.
The Danish tradition of trial projects

The contract management project should also be seen in the light of Denmark's tradition of using trials in the development of administrative policy.

In Denmark, administrative policy falls within the jurisdiction of the Ministry of Finance. This means that the Ministry of Finance co-ordinates major administrative policy initiatives and also undertakes the development of new administrative policy tools. In the relationship between the Ministry of Finance and the other ministries it is essential to be aware that Danish government ministries traditionally enjoy a high degree of autonomy. Thus, the individual ministers are the supreme policy administration heads in their domain. In practice this means that the decision to implement administrative policy initiatives is largely taken by the individual minister, with assistance and guidance from the Ministry of Finance.

This relationship between the Ministry of Finance and the other ministries has led to the development of a tradition for launching pilot projects or trials. Typically, the Ministry of Finance develops a project concept that is then tested by other relevant ministries to determine whether it functions as intended and to establish the concept's general practicality. The contract management pilot project is a part of this tradition.

The pilot projects typically run over a period of two to five years. This is seen as a sufficient time to evaluate whether or not the trial concept is practicable. The evaluation is done in co-operation between the Ministry of Finance and other involved ministries. It is often considered useful to get an independent evaluator to carry out the initial evaluation, that will be used as a reference in the final evaluation by the ministries. In other cases there is no formal evaluation, i.e. if a project is generally seen as a success there will be a tendency to integrate the trial concept into the practices of the ministries and the concept will become a regular management practice.

The criteria for success may vary between projects, the objectives can be improved efficiency, increased involvement by users or improved service quality. Success is generally seen not only from the viewpoint of the objectives set by the Ministry of Finance; the degree to which other involved ministries accept and value the trial concept is also important.

The Danish ministerial system

The structure of the Danish ministerial system is generally such that each ministry consists of a department and one or more subordinate agencies. The minister is formally both legally and politically responsible for the department and all its subordinate agencies. The department functions as the minister’s secretariat and also has a co-ordinating role vis-à-vis the subordinate agencies, that perform a very wide range of functions. These agency functions range from casework and research to the performance of highly specialised functions.

No hard and fast rules have been developed governing the control aspect of the relationship between a department and its agencies. In practice there is a wide variation between the ministries. However, a general characteristic of the relationship has been a tendency for the department to concentrate solely on its function as secretariat for the minister, whereby its control function vis-à-vis the agencies has been reduced. To obtain an increased focus on results, modification of the managerial role of the departments has therefore been a central challenge, both generally and in conjunction with the contract management project.
The Project

Launch of pilot project

In 1991 a decision was taken by the Cabinet to initiate a preliminary study into the potential for improving the efficiency of state agencies via a contract-type control system. The study was performed by the Ministry of Finance in collaboration with a firm of private consultants.

An analysis of 15 state agencies and departments revealed 31 barriers and other obstructions to improved efficiency. Significant barriers included absence of targets and result-related controls, planning problems due to budgetary uncertainty, inadequate freedom of action with regard to pay and appointments, investment problems, and difficulties confronting potential private sector collaboration. The large differences that existed between the problems of the individual agencies led to the conclusion that specific solutions should be created for specific problems. Or more precisely, that the problems of the individual agency should be made the point of departure and a unique contract established accordingly. The first contracts entered into force in January 1992, followed by a second wave in January 1993. A total of 13 agencies under seven ministries took part in the project.

The contract management concept - and contract contents

The principal aim of the contract management project was two-fold: to improve operational efficiency and to increase policy control potential.

Overall, the problems identified by the preliminary study pointed to the need for:

− increased freedom of action in the performance of tasks;
− budget guarantees based on multi-year agreements;
− establishment of clearer target- and result-related requirements.

Against this background the idea was that a result-based contract should be negotiated defining both the rights and the obligations of the agency. In other words, the contracts would be based on a something-for-something principle: the agency would promise to meet certain increased efficiency requirements formulated in terms of specific results, in return for which it would receive a certain freedom of action with regard to task performance plus a budget guarantee for a number of years.

All the project contracts had a four-year term. The contract consisted of a description of the tasks, rights and obligations that were to form the basis for task performance. In addition to the more specific performance targets, the obligations could include more process measures (i.e. measures of activities and initiates within an agency) and service quality oriented measures, e.g. new initiatives in respect of services or groups of users, requirements relating to personnel matters, or requirements for increased information provision to users and the public. The contract was also to contain a measurable operationalisation of the established targets and obligations, and a statement of the agency’s required reporting frequency. The reporting frequency for all the contracts covered by the project was set at once per year. The contract was not legally binding in the normal sense as it could, in principle, be revoked by the minister.

The idea was that the subsequent reporting by the agency should both pressure the agency to make efficiency improvements and open the way for increased insight by the department into the agency’s work. No sanctions for failure to reach performance targets were specified. However there has not been a
need for sanctions as all agencies have achieved the targets defined in the contracts. One of the main objectives of the project was also to improve co-ordination and connect together the department’s specialist and financial management functions.

Examples of performance targets and special conditions provided in performance contracts are shown in the Annex.

Selection of project participants

The agencies suitable for project participation were mainly chosen by the responsible ministries. The main selection criteria were that:

- an agency reorganisation process was already under way;
- an agency was already functioning effectively;
- an agency was motivated for change;
- there was a clear need for efficiency improvements;
- there already existed studies or action plans that could be used as a starting point for contract negotiations.

Project findings

Results

In December 1995 the Ministry of Finance published a report containing the provisional findings of the contract management project.

The findings showed, firstly, that a substantial number of the barriers to efficient agency working had fallen away. In general, the participating agencies were seen to have fulfilled the targets and result-related requirements established in the contracts. Thus, the result-based contracts did lead to increased agency performance in terms of improved efficiency and service quality. To a large extent the contracts acted as a hub for agency reorganisation processes and efficiency activities, such as development of internal management and control tools and targeted developmental activities in the personnel domain. In addition, contract establishment often served to increase personnel awareness of work place objectives. In many cases agency performance was found to be better than the contract specifications. This might suggest that in some cases the requirements were set too low.

Secondly, the findings showed that provision of information to the public and users was strengthened. The contracts generally focused strongly on level of service, case handling times, information, and so on. The views and satisfaction of the public and users were directly incorporated in the contracts, partly in conjunction with the need analyses on which the required results were based, and partly in conjunction with requirements for user surveys.

Thirdly, the contract management project generally resulted in an improved dialogue between agencies and departments. The intensive contract negotiations led to the creation of a network of contacts between the parties at all levels. The departments gained improved insight into the workings of the agencies, that led to a more holistic, all-round dialogue rather than the traditional ad-hoc dialogue focusing
on detailed issues. In addition, the reporting provided improved documentation of agency activities, something that in turn improved the quality of communication.

Fourthly, contract management proved to be a useful tool for changing agency priorities in line with changed external demands -- not least new policy requirements. The potential for policy control was improved as a result of the improved information platform. At the same time better integration of programme management and financial management was achieved, something that was very largely lacking in the normal one-year budget negotiations. However, there was criticism in the evaluation that the excellent strategic control potential of contract management was not adequately used by ministers and departments. Nevertheless, positive results were achieved, even with the limited departmental strategy participation to date.

Finally, the project showed that it was possible to develop the necessary methodology for measuring and documenting even highly complex activities. As well as the conventional quantitative measures (efficiency, savings), success was achieved in developing systems for measuring the quality of case handling. In addition, two agencies succeeded in establishing operationalised requirements for quality improvements. Examples of “soft” measuring tools contained in the contracts include user surveys and image analyses. This was one of the most important results in the project.

Negotiations and participant attitudes

The agencies themselves had a major influence on the establishment of the result-related requirements and target setting. The Ministry of Finance played an important part in the negotiations with the agencies on the elaboration of the first contracts, so important that the Ministry of Finance - rather than the individual responsible department - was often regarded as the opposite party in the negotiations. The role played by the Ministry of Finance was necessary to set the project in motion and to arrive at a form of standard model for the contracts. In the future it is intended that the Ministry's role in the negotiations will be taken over by the individual responsible departments as a means of increasing their feeling of responsibility towards the contract. The sole role of the Ministry of Finance will be to advise and ensure continued development of the concept. This will also increase the departments’ strategic action potential and their ability to reflect their relevant policy goals in the contracts.

At the time the first contracts were established, the departments were sceptical about the project. This was partly because the budget guarantee given to the agencies could be seen as a limitation on departmental freedom of action. For this reason, the departments left the negotiations largely to the Ministry of Finance. However, by the time of the second wave of contracts, the departments had become more alert to the potential of contract management, and the selection of agencies for new contracts was based on a deliberate choice of control methodology.

Most of the agencies saw the contract management project as a means of setting the agenda for their future development, and thus often acted as initiative takers and formulators in the framing of contracts. However, the agencies showed strikingly little interest in additional autonomy, which is surprising in view of the barriers to efficiency that emerged in the preliminary study. This can partially be explained by the barriers being more formal than real, and some barriers were removed more generally for all agencies during the contract period. By contrast, substantial weight was placed on the budget guarantee, principally with regard to the possibilities opened up for long-term planning. Thus, one or two agencies commented in the evaluation that if the budget guarantee were removed there would be nothing left to negotiate about. Within the ranks of the management, there was particular satisfaction at being able
to point to external pressure when called upon to implement changes and efficiency measures. The agencies were generally satisfied with the improved dialogue with the department.

**Problems**

The crucial factor in contract formulation is to establish operationalised result-related requirements that at the same time cover the full range of agency functions. Already from this it is clear that contract management should principally be applied to agencies for which targets can be established for all major tasks. An agency will focus on the targets stated in the contract, and if there are components in an agency's functions that cannot be measured, and that are therefore not included within one of the performance criteria in the contract, the function concerned will probably be assigned reduced priority by the agency. In conjunction with the contract management project, particular attention was drawn to the danger of excessive weight being placed on quantitative result-related targets at the expense of process and service quality targets related to, for example, internal backup functions, public and user information, and personnel policy goals. For the same reason, the selection of agencies for the project was often based on the pre-existence of a reasonable overall picture -- and ideally an overall action plan -- for the agency. Without this pre-preparation it would be difficult to identify all the key factors and thus enable a balanced list of result-related requirements to be established.

As mentioned, one of the most significant results of the project was the successful development of tools for measuring complex activities. However, there still exist activities that are not directly measurable, something that can lead to distortion of the incentive structure. Similarly there are cases where it is difficult to form a reasonable picture of the agency's task profile. Generally, however, the study found very few tendencies towards distortion in agency development. This was due partly to the success in the development of measuring tools, and partly to the fact that Danish agencies normally have a very strong specialist profile in which specialist competence and identity prevent reduced priority from being given to areas difficult to measure. However, the avoidance of distortion in the incentive structure was also due to the fact that agency selection for the project was specifically based on the pre-existence of an overall picture and documentation for the agency's task profile.

The process of contract establishment is extremely resource-intensive. Contract formulation takes a long time, but the subsequent resource burden is reduced by the fact that ongoing control is simplified. However, some areas of public-sector activity are characterised by the presence of many agencies with uniform tasks. In such cases the project showed that negotiating a customised contract for each individual agency is too costly a solution.

**Future Use of Contract Management -- Deliberations and Problems**

Today, result-based contracts are operational in 20 agencies and more are on the way. Most of the agencies in the pilot project wish to continue with contract management. However, what will happen when the individual departments take over negotiations is an open question. In the light of the evaluation and experiences with the contract management project, various changes have been introduced into the concept and new changes are on the way.

Changes currently being considered or already introduced include the following:

Firstly, consideration is being given to whether the budget guarantee should necessarily form part of future contracts. The Ministry of Finance has clearly stated that it will not in future provide a
guarantee and that any such budget guarantee must be provided by the department within its funding allocation. The practice hitherto has obscured the departments’ control responsibility. Therefore, the question is whether the departments wish to provide their own guarantee.

The guarantee is a useful bargaining tool vis-à-vis the agencies, and experience has shown that they accord it considerable value. The problem is that the budget guarantee reduces the number of action alternatives for departments and ministers. The budget guarantee may therefore become an encumbrance on the future use of contract management. In addition to this, it can be argued that result-related requirements are a wholly natural imposition, and that agencies should not, therefore, be given something "extra" for taking on such obligations. Whether the budget guarantee should be included in future contracts will be up to the individual responsible ministry to decide. Consideration could be given to use of a more dynamic, rolling form of contract within which the budget guarantee is not an indispensable component and within which new guarantee periods could be introduced as existing periods expire. With this system, renegotiation will arise at regular intervals, providing ongoing recourse to sanctions on the agencies. Another possibility is to make the budget framework dependent on the result targets. However, this would require a substantial knowledge of the relationship between activity and cost, and this information may be difficult for the department to obtain.

Secondly, consideration should be given to a possible widening of the policy dimension in the contracts. The objective of using contract management as a means of strategic control has not been adequately fulfilled. The rather unobtrusive role played by the departments during the project should be changed, and the minister should adopt a more active role. The departments are now alive to the strategic control potential offered by contract management. At the same time it is important that the departments should involve themselves actively if a qualified counterbalance to the agencies is to be established at the negotiating table.

Thirdly, consideration must be given to improving the agency annual reporting. This procedure was in several cases not satisfactory. In future, increased weight should be placed on the reporting function as good reporting is crucial to the agency's continued implementation of the result-related requirements in the contract, and can also serve as a continuous reservoir of experience and a learning process. The work surrounding the reporting system will in future be integrated with another Danish project focusing on annual reporting. The purpose is to ensure that all state agencies with operations-related and outward-oriented functions submit annual reports in a form capable of providing increased information on results.

Fourthly, consideration should be given to the contexts in which contract management is well suited. Experience has shown that contract management is useful for supporting operations- and reorganisation-related tasks and can usefully supplement other control methods. However, it should in future primarily be used by ministries that administer large numbers of agencies with different types of tasks. Contract management is less suited for agencies where it is not possible to give a relatively precise and comprehensive description of relevant work tasks and development needs. In order to expand existing knowledge of where contract management can be suitably applied, it is planned to establish a contract management experience network where guidance can be obtained when new contracts are to be established.
ANNEX

Examples of Targets and Special Conditions of the Danish Free Agencies

The following mentions the most important elements in the listed free agency contracts:

Risø National Laboratory (Ministry of Research) 1994-97

The Ministry of Research seeks in this contract to pioneer the application of performance management and results-based contracts to research institutions. Risø, a research institute that seeks to support industrial and commercial development, is described as a good pilot candidate because of its considerable size, and its ability to adapt to changing conditions. Risø proposes to be evaluated by a series of objectives and targets established by its Board of Governors, known as “RISØ 2000.”

Risø's overall objective is to create new opportunities for development in energy technology and energy planning; environmental aspects of energy, and industrial production; and on materials and measuring techniques for industry. Research on these areas and nuclear matters is distributed over 11 programme areas, for which targets have been established for the next 5-10 years. The 11 programme areas are in turn subdivided into 30 subprograms with 3-5 year plans.

Results are evaluated at the end of each year, partly on the basis of achievement of milestones defined in Risø’s research programme and partly based on performance on three sets of quantitative indicators: activity indicators concerning personnel resources; communication indicators; and network indicators concerning relations with Danish and international research. Risø will provide annual reports on its results and a final report upon completion of the project.

State and University Library (Ministry of Cultural Affairs) 1993-96

The contract lists the overall objective as being to extend the library's position as a central distributor of information -- the chief means being organisational development and use of information technology.

Key targets include:

- Improve efficiency through meeting a 10 per cent increase in the number of loans;
- Increasing revenue by 2.5 per cent per year;
- Retroconvert the card catalogue (1 million records);
- Formulate a policy on withdrawal of materials and increase withdrawals by 95,000;
- Introduce a plan for quality control of services involving different user groups;
- Carry out user surveys in 1994 and 1996 comparable to the 1988 survey; and
- Continue organisational development through evaluation of management with participation of external consultants.

The National Museum (Ministry of Cultural Affairs) 1992-95

The contract includes a declaration of aims, and then a series of measurable targets, including museum attendance increases, a change from a financial deficit to a profit, and an increase in the number
of articles published in internationally recognised periodicals. Each year, until completion of the contract at the end of 1995, the Museum will submit a report on progress towards meeting the agreed targets.

**Court of Taxation Appeals (Ministry of Taxation) 1993-96**

The court processes appeals of tax assessments and real estate valuations from the local tax authorities.

**Targets**

- reduce the average time a case is processed from 14 months in 1992, to 9 months in 1995;
- deal with all complaints related to the 19th general valuation of real estate before end 1996;
- increase productivity 20 per cent by 1996.

**Special conditions**

- long term budget arrangement to be provided;
- additional manpower to be provided;
- performance-related pay schemes to be introduced.

**National Board of Industrial Injuries (Ministry of Social Affairs) 1993-96**

The Board of Industrial Injuries processes applications for compensation for industrial injuries.

**Target:**

- reduce the proportion of cases where the processing time target is exceeded from 23 per cent in 1992 to 19 per cent in 1996;
- increase productivity from approx. 156 standard cases per employee in 1993 to approx. 162 per employee in 1996;
- reduce the number of appeals from 18 per cent to 14 per cent by 1996.

**Special conditions**

- long term budget guarantee to be provided;
- possibility of working at home to be introduced;
- special days off in case of children's illness to be provided;
- performance-related pay schemes to be introduced.

**Medicine division of the National Board of Health (Ministry of Health) 1993-96**

The Medicine division is responsible for licensing drugs prior to their marketing, for inspecting products on the market and for inspecting chemists.

**Targets**

- increase yearly inspections of producers by 20 per cent and increase in yearly inspections of chemists by 75 per cent before the end of the contract period;
- increase from 50 per cent in 1992 to 80 per cent in 1996 of the amount of applications for parallel importation that are processed in less than four months;
- ISO-9000 certification.
COMMENTARY
by Professor Anne Drumaux
Ecole de commerce Solvay - Université Libre de Bruxelles

The Contract Management Project in Denmark

Planning or “management” contracts\textsuperscript{17} traditionally aim to improve the performance of public enterprises or bodies. The idea of contractual autonomy, formalised at a very early stage in France\textsuperscript{18}, has developed through different experiments, all of which originated in the dissatisfaction with traditional budgetary control mechanisms and the need to establish performance control of efficiency (for example productivity) and effectiveness (quantity and quality of services). Of particular note, in relation to the OECD Member countries\textsuperscript{19}, are the US Government Performance and Results Acts, the contracts in Germany and the Netherlands between regional authorities and railway undertakings, France’s planning and objectives contracts for the railways, electricity and some commercial and industrial public bodies (EPICs), and Belgium’s management contracts for the railways, telecommunications and low-cost housing. In the United Kingdom, the use of contracts is part of the broader initiative to introduce market mechanisms into public services ( contracting out, contracting in)\textsuperscript{20}. It is clearly the failure of traditional methods of regulation, whether of public bodies or private operators, that lies behind the development of contractualism\textsuperscript{21}. In this sense, the Danish experience fits in with the tradition of reforms aimed at contractualism. The needs identified are the same as in contracting experiments, that is, to:

- increase freedom of action in order to improve performance;
- establish budget guarantees on a multi-year basis so as to remove the constraints imposed by the principle of yearly budgeting;

\textsuperscript{17} The different terms used here synonymously sometimes arise from somewhat different traditions which cannot be discussed here in detail. It is also relevant to mention the specific terminology of the World Bank which uses the term performance contracts for contracts between public enterprises and their responsible authorities, keeping the term management contract for contracts between a public authority and a private operator. See World Bank (1994), \textit{Infrastructure for Development}, New York, Oxford University Press.


\textsuperscript{19} Projects have also been implemented in India and Latin America, see World Bank Report (1994).


– define objectives as clear targets, measured in terms of expected results, so as to be able to move from the control of compliance and legality to a full assessment of performance.

Having said that, the Danish experiment is totally original in a number of respects:

– the role played by the Ministry of Finance in elaborating the contracts is altogether quite unusual: project promoter at the first stage of experimentation with contracting, advisor or consultant at the development stage. This change of role is interesting and it will be addressed below when the incentive aspects of contracting are discussed;

– the project seems to have covered a very wide range of activities in different ministries (research, cultural affairs, taxation, social affairs, health, etc.) and to have involved administrative agencies, that is a rarer occurrence in other countries. As a result, the experiment led naturally to discussion about measuring the performance of complex activities. It is acknowledged that administrative activities are generally intangible services whose characteristics need to be detailed very carefully to understand what constitutes their quality. The complexity of the activity sometimes makes it difficult to define administrative output and identify the direct or indirect beneficiaries;

– the concern with assessing the reform (over and above the Ministry of Finance’s role already mentioned) as it developed is indicative of an exemplary administrative process.

The conceptual framework for analysing contractualism practices

Before commenting on the Danish reform, the contract analysis framework needs to be explained. Different theoretical tools are available: the contract model with the agency theory, the transaction model with the transaction cost theory, the convention model with the convention theory. The perspectives of these three approaches differ and will be described very briefly below.

Put most simply, agency theory\(^\text{22}\) attempts to resolve co-ordination between a principal and an agent, and is focused on resolving information imbalances through the use of incentive mechanisms in contracting. By focusing on the incentive aspect, the motivating factors in actual contracts can be analysed, for example, as can the quality of the control exercised by a responsible authority.

The aim behind transaction cost theory\(^\text{23}\) is to explain the organisational structure and control procedures that co-ordinate agents at the lowest transaction cost. This is a tool that can be used, for example, to compare contracting-out and contracting-in procedures (i.e. “buy” or “make”).

Conventions theory\(^\text{24}\) looks at the co-ordination problem from the point of view that motivations are based not on the pursuit of agents’ personal interests, but on socially internalised mechanisms. A convention is a collective agreement, tacit or explicit, that specifies how to behave in a recurring situation and so allows agents, once the convention has been established, to co-ordinate their actions merely by


foresight. Organisations thus establish conventions that distinguish between what belongs to individuals or operational units and to their independent action, and what is in the domain of the hierarchy and line control.

Which of these three paradigms to use (contract, transaction, convention) depends on what is being analysed: the contract model for effectiveness of control, the transaction model for contracting out decisions and the convention model for work relations.

The contract model will be used in this analysis to illustrate the information and control problems facing the responsible authority. Since the contracts analysed are geared towards the production of outputs, this approach is totally justified inasmuch as, to manage these contracts, the responsible authority has to evaluate and measure outputs. Therefore, the quality of information on these outputs is a key issue.

A contract between a responsible authority (the Ministry of Finance in this case, at least during the first stage in the reform) and a public enterprise or agency can therefore be presented as a relationship between a principal and an agent, binding them and committing them to doing or not doing a number of things.

Each of the parties can pursue different objectives. There would, in fact, be little point in negotiating a contract if the interests of the parties were exactly the same. What is more, the principal and the agent are not necessarily in possession of the same information when the contract is negotiated.

This information imbalance stems primarily from the fact that the principal cannot possibly know the agent’s productivity level or quality standards. This situation is called adverse selection as it is the agent who has the best information on the production conditions (productivity, quality), while the principal in fact knows only what the “other side” divulges.

The information imbalance also stems from the fact that the principal is not necessarily in a position effectively to check the agent’s level of effort (a situation referred to as moral hazard) and cannot therefore reward the agent according to the intensity of the effort.

The theory of the principal-agent relationship should be borne in mind when discussing the quality of contractualism reforms in government.

Comments on the Danish reform

The contract philosophy in the Danish experiment is relatively typical compared with previous experiments. For the reasons stated above, the comments will discuss the contract as an instrument of information and control for oversight within the framework of an agency theory approach.

The comments will chiefly focus on the following issues:

− the idea of commitment in contractualism procedures;
− budget guarantees;
− incentives;
− transfer of experience; and
− standardisation of procedures.
**Commitment**

The absence of sanctions in contracts is a key element of discussion. While the information imbalance between agency and principal is a reality, widely confirmed by research into administrative behaviour, it should be acknowledged that the absence of sanctions may be seen as standing in the way of fully effective reform. Target setting and the budget guarantee will be discussed below. The point to note at this juncture is that sanctions (positive or negative) are a way to reinforce the agency’s commitment to achieve a target, particularly when the context shifts from a stimulating pilot project to a reform establishing a recurring practice.

To make the agency’s budget framework dependent on results is one possible stimulus, but it is not the only solution. Generally, the sanction can be positive (to obtain a benefit or not, as the case may be) or negative (to incur a penalty in the event of certain dysfunctions). The use of one or another of these solutions depends upon the seriousness attached to failure to meet a target. Penalties should clearly be reserved for flagrant and proven cases of, for instance, non-compliance with quality standards.

The fact that the contract can be revoked by the minister makes the contracting procedure shaky if there is no appeals procedure. The possibility that the contract can be revoked may be interpreted by the agency as a sign of minimum commitment by the principal, that would inevitably have an adverse effect on the agency’s own commitment in return.

The use of a third party is a conventional method of offering some security. The question arises as to why the Ministry of Finance might not play this role in the subsequent stage in the reform, considering its culture of experimentation and hence expertise in this field, as well as the structure of the Danish ministerial system.

**The budget guarantee**

The budget guarantee is a key issue. To what extent is achieving results a natural obligation that need not be “rewarded”? The problem touches again on the information imbalance. The agency is the best observer of its actions. It may be pursuing its own particular goals that are not necessarily in step with the principal’s objectives, that are to achieve a particular result. Moreover, all organisational theories highlight these strategic behaviour patterns that are not specific to the public sector. The information imbalance provides a built-in information advantage when there is overconsumption of resources in relation to results, or a deterioration in quality when cost constraints are tight. The fact that public organisations operate in the public interest does not protect them from these strategic behaviour patterns.

Therefore, the conclusion is that, if the consequences of this are to be avoided, mechanisms are needed that will prompt the agency to behave in the desired way of its own accord and not to rely solely on instructions, whether in the form of a mere reminder of mission statement obligations or a performance standard imposed from above.

If the budget guarantee becomes a departmental responsibility, budgets for other departmental activities could be reduced. If departments and the minister responsible are to make optimum use of the strategic control tool of contracting, there necessarily has to be a commitment, not just symbolic in the spirit of reform but also budgetary, failing that the reform could lose part of its purpose. By granting a certain budget guarantee to agencies pursuing objectives considered to be priorities, the department

---

25 The issue has been well documented by the Public Choice School.
becomes directly involved in achieving results. It can be argued that the objectives in current contracts are in fact connected with existing functions rather than new ones: improving the quality of library services or reducing the time taken by the tax authorities to process cases, are not new functions.

At the preliminary stage of the transfer to departments, there could be negotiations between the Ministry of Finance and departments to encourage them to select contractual forms that offer more incentives. Neither is it sure that the Ministry of Finance has to be the only consultant: the use of outside consultants could help departments to see the benefits of reform objectively.

The relative lack of interest shown by agencies in greater freedom of action, and their concern for a budget guarantee from a long-term standpoint, are factors to be taken into consideration for the future of the reform. It is possible that freedom of action is not the priority goal for agencies, either as stated in the report because the actual barriers have fallen away, or because independence is not a sufficient incentive for some agencies. Whereas, that does not seem to be the case for the multi-year budget guarantee.

It must be recognised that the search for a budget guarantee is consistent with the idea of contractualism in that it reduces the constraints inherent in the principle of yearly budgeting. The budget guarantee acts as a sign of the principal’s commitment and is therefore vital, but there is nothing to say that it should not be adjusted if necessary. The agencies’ commitment in return also needs to be ensured. The solution probably lies in finding mechanisms that offer more incentives.

**Incentives**

While contracting has clearly facilitated the reorganisation of agencies by promoting efficiency and effectiveness (intrinsic quality of services, starting with a more precise definition of complex administrative activities and culminating in concrete indicators of delivery and quality, managerial reform at personnel management and financial levels, user information, quality of communication between the agency and the Ministry of Finance, etc.), the incentive aspect of contracts is still open to question. The absence of sanctions comes down to the fundamental problem of the information imbalance between principal and agent. As the report notes, the fact that there was no need for sanctions because the targets were achieved suggests that these targets may have been set too low. This is at the heart of what the theory terms information imbalance: “The agencies themselves have had a major influence on the establishment of the result-related requirements and target setting”\(^\text{26}\). The dynamics were positive, but there is nothing to suggest that the process could not have been taken further: productivity and efficiency could inevitably have been improved.

It would be useful to make the contract more motivating, encouraging the agency to exceed the targets. One possible way to do this is to abandon the single contract and introduce a menu of contracts. The idea is simple: when the agency is confronted with a choice between several contracts, it will be required to make a choice that will reveal some of the hidden information. To put it more simply, imagine, on the one hand, a contract offering a minimum objective combined with an unattractive cost repayment system or the prospect of a minimum budget guarantee and, on the other hand, a contract setting a maximum objective combined with a higher payment (better cost repayment, but also longer

---

\(^{26}\) This finding, highlighted by the report, is common to many analyses of contractualism practices, see Haut Conseil du Secteur Public (1992), Les entreprises de service public, rapport 1990, Paris, La Documentation française. The finding was similar for Belgacom, the Belgian telephone enterprise, documented in Debande O, Drumaux A. (1994), Regulating Infrastructure in Europe: contractual approaches, Background Paper for the 1994 World Development Report, Washington, The World Bank.
Another possibility is the subsequent development of the reform, particularly where several comparable agencies might be engaged in parallel contracting. This would involve multilateral contracts with agencies made interdependent: agency x would be rewarded (by whatever means) both in terms of the degree of achievement of its objective, and also according to its success compared with other agencies.27

The proposal for “rolling contracts” introducing new guarantee periods when the contract expires, seems to be a strictly technical issue about how to transfer existing contracts to departments. By way of contrast, the budget guarantee is a policy issue as it concerns the department’s commitment to the reform. The possibility of making the budget framework dependent on results is one possible way to strengthen the agency’s commitment. The objection regarding the information unavailable on the link between activity and cost-generated is not totally justified: it is true that there is an information imbalance against the principal on this point; using an appropriate technique of incentive-based contracting, such as contract menus or multilateral contracts, it is possible to overcome the barrier and obtain some of the hidden information. But is there a danger of budget overruns with a system of “rewards”?

In practice, it is possible to establish a link, in some cases a partial link not covering the entire budgetary framework, between inputs and results if a system of positive sanctions is adopted. Contract menus and multilateral contracts can normally smooth out cost overruns: the improvement in service efficiency and effectiveness should not be measured solely in relation to the cost of more motivating contracts, but should also be measured in terms of the gains deriving from the achievement of targets. Examples might include the increase in library revenue, museum attendance and productivity.

**Transfer of experience**

The little amount of interest shown by ministerial departments in using management contracts as a strategic control tool is due to the fact that the pilot stage was managed by another principal, the Ministry of Finance. In this sense, the proposal to transfer the role of principal to ministerial departments is likely to strengthen their commitment. However, there is still a risk that the departments’ sceptical attitude toward the reform, as a limitation on their freedom of action, could water down that commitment. It is therefore sensible for the Ministry of Finance to retain the role of advisor and promoter of the spirit of reform.

The forms of the Ministry of Finance’s involvement need to be clarified. As mentioned, negotiating contracts is very time-consuming and hence resource-intensive. Therefore, the Ministry of Finance has developed an “engineering” approach: preconditions for contracting (suitable types of activities, types of indicators), negotiation programme, forms of contracts, etc. What remains now is to organise the effective transfer of its administrative management capabilities to departments. As stated above, another possibility is to give the Ministry of Finance the additional role of arbitrating in the event

---

of conflict over a contract: all the experiences with contracting have shown that the commitment of parties to the contract needs to be sustained.

The first experimental contracts have made it possible to define the types of agency that could operate under a form of contracting (ability to define measurable targets for all the agency’s functions and tasks, precautions so as not to give excessive weight to quantitative targets, identification of the agency’s action plan).

This lends extra weight to the proposal that the Ministry of Finance should retain an advisory function. Moreover, while the quality targets, whether concerning service quality, internal backup functions, user information or personnel policy, are naturally qualitative, they are not necessarily unquantifiable: frequency of breaks in service, dysfunctions, case handling time, percentage of complaints, internal and external satisfaction surveys, etc.

**Standardisation of procedures**

Contractualism is costly in terms of time and, therefore, resources. The possibility of standardising contracts is a solution in the case of agencies performing the same kinds of tasks. It may be more difficult to persuade agencies to agree to negotiate standard contracts. It should also be said that the benefits of better communications and mutual information during the negotiation process, highlighted in the report, could be diminished by contract standardisation.

A possible compromise solution would be to define a “common core” for contracts containing the objectives common to the agencies, with the possibility of adding a limited number of “made-to-measure” targets. The advantage of this solution is that it reduces costs while leaving room for negotiations -- without which a contract is not a contract. As mentioned above, multilateral contracts are a variant of this solution: these are contracts that are partly the same and contain payment mechanisms that bind agencies to each other. Not only does this save on the transaction costs of negotiations, but it also limits the built-in information advantage.

3. **Some Key Elements for Contracting Practices**

First and foremost, to be aware of the conflict between objectives aimed at improving productivity or achieving a quality target, and the aim of minimising the built-in advantage to the agent. The built-in advantage arises from a non-optimal commitment by the agent to achieve the objective, that allows the agent to derive a real surplus through a resource allocation that is too high for efficiency or effectiveness. Harsh treatment pays no dividends. Even when the contract includes a mechanism that rewards increased productivity, for example, an incentive mechanisms needs to be included based on the difference between predicted and achieved productivity gains.

---

28 This solution was adopted for low-cost housing companies in the Brussels-Capital Region project mentioned above.
A good contract should reveal hidden information. Instead of a single contract, the principal may propose a menu of contracts involving different levels of coverage of costs and commitments. By choosing one of these contracts, the agent in fact reveals the level of commitment that the agent actually can and wants to envisage. The multilateral contract is another solution through which information is revealed by comparing between different agencies performing similar tasks.

Adverse selection and moral hazard inevitably lead to a loss of control by the principal. To reduce the information imbalance, the principal may supplement the control system with external audit procedures. Likewise, parallel contracting with different agents can be an appreciable source of information (yardstick competition).

In order to strengthen the level of commitment, it is preferable to limit the possibility of renegotiating active contracts. To provide contractual security, a third-party authority should be nominated beforehand to intervene in the event of conflict.