EVALUATION KNOWLEDGE FOR STRATEGIC BUDGETING

Xavier Ballart and Eduardo Zapico

INDEX

I. INTRODUCTION

Rational budget reform has traditionally claimed the instrumental linkage between resource allocation and evaluation. Despite limited advances, this is still one the aims of many current budget reform initiatives. This chapter deals with a number of issues related to the links between budgeting and new approaches to evaluation, taking into account the complexity of the public management context. Adapting budgeting to respond to the growing uncertainty and interdependence requires new criteria of success and evaluation as the basic framework of reference for management performance assessment. New approaches to accountability can also work as an incentive to encourage interactive learning.

Much of the recent thinking and research in the area of trust has been informed by work that emphasises the role of human capital and social capital (OECD, 2001b). Human capital focuses on the importance of a wide range of individual capacities, including informal learning, while social capital is a property of groups, communities and organisations that emphasises the role of shared norms and values as well as social behaviour. Very close to this notion is the role of trust as a crucial factor for social co-operation.

In particular, we would like to explore the new roles of central budget office (BO) in strategic resource allocation and the opportunities offered by new approaches to performance measurement for evaluative knowledge to become part of the budget process. We have reviewed potential opportunities for strategic leadership under uncertainty along the following lines:

- Developing new criteria of success to assess performance
- Applying alternative approaches to evaluation
- Building trust for evaluative knowledge to be effectively used
- Redesigning positive oriented accountability systems

Our main concern is to identify and assess new arrangements that facilitate strategic budgeting using evaluative knowledge. The creation of formal structures and institutions to facilitate

technical evaluation studies in the framework of the budget procedure will not suffice. Neither any simple formula based on well-identified evaluation models. Many conditions and factors are required to foster an environment conducive to the production of useful information, mutual exchange and widespread organisational engagement. We envision evaluation as a space of knowledge creation (not necessarily by evaluators or budgeters) and BO as having a role in gathering information, managing knowledge, integrating it into existing knowledge, and promoting social learning by sharing applied knowledge to create value for stakeholders and the society at large.

II. STRATEGIC BUDGETING AND EVALUATION UNDER UNCERTAINTY

Strategic Leadership from Central Budget Office

The interest of public managers in using evaluation and public expenditure management tools may be quite different from the interest of BO. Any agency or manager should develop its own budget management and evaluation system without the centre necessarily playing a role, beyond legal compliance and aggregate spending discipline. However, if evaluation and budget reform is to be part of an overall initiative for shifting from a traditional direct type of control to a more flexible type of spending management control, the need for central leadership seems to be clear.

The specific role of a central budget unit and the way it exercises its leadership for decentralisation under a context of uncertainty need to be discussed. The relevant question is: what does central steering mean under a decentralised and uncertain policy context? This question suggests the need to change and adapt performance management, evaluation and other related functions such as contract management and accountability. (Office of Management and Budget, 1995).

With decentralisation, changes in the responsibilities of the central unit are not just quantitative but also, and more importantly, qualitative. Current public sector reforms such as contract management, budgeting for results and decentralisation in general demand improvement in the management capabilities in both decentralised agencies and in central units. The former needs to

undertake better evaluation and reporting. The latter needs understand evaluation and also to be able to provide leadership for the introduction, development and co-ordination of a government-wide decentralised management system among interdependent organisations, including evaluation. If, during the process of decentralisation, agencies resist improving their evaluation, the centre must actively play a leadership role in promoting evaluation, so that the government as a whole can learn from its experiences and each organisation does not have to begin from scratch.

Some might argue that after decentralisation the new role of the centre is more limited as it would be circumscribed to co-ordination and post-evaluation (Shick 1990, p. 33). In this study we argue the opposite, that is, that the centre needs to play a qualitatively 'stronger' role after decentralisation. We suggest that after decentralisation the performance of the government-wide management system, in a context of complexity, heavily depends on adequate strategic leadership from the centre. The central budget unit has to pay special attention to increasing pressures towards fragmentation in a system struggling to work under a new type of accountability based on strategic direction and demonstrating performance (Mayne, 1999).

Independent budget behaviour and competition for resources might follow decentralisation if there is no special effort from central units to improve co-ordination (Zapico, 1999). Central budget's evaluation role is essential for the effective performance and management of the whole system. Furthermore, budgeting and evaluation are not just technical exercises. After decentralisation in an unstructured or complex environment, it is necessary to increase the capacity of the centre for budget flexibility and strategic leadership. This implies, among other factors: overall priority-resetting; negotiating and designing new rules for the budget game; building proper channels of communication and developing institutional capacity for co-ordination and conflict resolution. It also means applying new criteria in evaluation and accountability in order to ensure the integrated functioning of the decentralised parts of the government.

Interest and support for performance budget reform from senior officials and central departments may appear and develop if the performance measures chosen for evaluation are considered

relevant and credible. This suggests the need for redesigning the performance criteria. New types of performance measurement (the 3Ds: capacity for Diagnosis, Design and Development), which are discussed later can help to define further the strategic evaluation role of BO under conditions of uncertainty.

Furthermore, new approaches to evaluation and some of its implementation requirements are also needed and suggested to deal with the obstacles blocking strategic leadership capacity from BO.

Figure 1. Strategic Budgeting and Evaluation Under Uncertainty

Unstructured Policy Context	Related Problems for Resource Allocation	New Responses and Approaches To Evaluation
Uncertainty: unknown cause-effect relationships; difficult measurement or data collection	Limited scope of traditional budget and performance indicators	<u>Strategic Leadership from Central Budget Office</u>
	Unintended costs and benefits	1. New Criteria of Success: From 3Es to 3Ds
Interdependence: interaction among units, organisations, policies and constituents	Ambiguity of goal definition	2. Approaches to evaluation under uncertainty
	inability to control external factors	
		3. Building trust for evaluation knowledge utilisation
	lack of co-ordination among interdependent actors and budget programs	4. Positive Oriented Accountability Systems
Diversity: multiple and conflicting values and interests		
	resistance to control and evaluation	
Instability: changing goals, policies and environment		
	irrelevance of goals and spending policies	

As shown in Fig 1. the more the public management context is unstructured, -that is, characterised by the factors of uncertainty, instability, diversity and interdependence-, the more the need for new criteria of success and evaluation paradigm to be developed. Related problems for resource allocation have been identified, such as limited scope of performance indicators, unintended effects, ambiguous goal definition, controlling external factors, lack of co-ordination and resistance to control.

Therefore a core condition for active and strategic budget leadership in an unstructured context from a central budget unit is having capacity to confront diversity, interdependence, uncertainty, and instability. This context demands new responses to public budgeting and evaluation, including new performance criteria and evaluation approaches.

New Criteria of Success: From 3Es to 3Ds.

Major obstacles for proper budgeting, such as contextual and policy instability, participants interdependence, diversity of interests and, in general, unknown program effects can often best be addressed by rethinking the nature, scope and approach of evaluation. We need to reinvent evaluation and redefine the criteria of successful performance to match today's organisational realities (unstructured policy context demanding decentralisation). Success with both evaluation and budget reforms requires responding to several sources of uncertainty.

The definition of "sound" performance used for evaluation should reflect the adapting responses to uncertainty. According to Mayne, "a well performing public program or service is one that is providing, in the most cost-effective manner, intended results and benefits that continue to be relevant, without causing undue unintended effects" (Mayne, 1999). This definition and the need for *continued relevance* has important implications for the evaluation of policy performance and the models chosen to improve public management and budgeting.

In this respect, it is important to firstly clarify how uncertainty affects and influences public managers' performance.

In a context of uncertainty about outcomes of programs and/or policy instability, the ability to properly *redefine* goals and intervention modes is a precondition of success for public managers. The capacity and skills of managers to be economic, efficient, and effective - the 3Es - in achieving goals may be insufficient if the goals are changing or outcomes cannot be readily anticipated. In fact, the greater the organisational uncertainty, the more distorted will be the performance measurement system in use (Carter ,1991) and the less relevant is the use of the 3Es model for evaluation and assessment. This is because the model assumes the availability of precisely pre-defined and stable objectives and a well known (or at least widely accepted) cause-effect relationship between organisational activities and outcomes.

It is widely believed that some policies are not susceptible to evaluation under a traditional 3E approach. This is, for instance, the case of community development initiatives. Community development practitioners (Hugues and Traynor, 2000) frequently point out that their work cannot be evaluated fully, because, the evaluators do not fully understand the process of community development. In this respect, the principal issues that tend to be raised are:

1. Time scale (how long does it take to bring about social change in a community)
2. Causality (how can change or growth be clearly attributed to the social intervention)
3. Diverse views on success (how will a credible data-set be assembled whereby to appraise outcomes when community members, community development staff and planners each have a different concept of what constitutes success)
4. Focus on process (sometimes seen as an end product in its own right)
5. Tendency to list activities on the assumption that such processes are likely to contribute to the well being of the community. In other words, the setting up of structures (women's centres, health centres, crèche facilities.) and processes (education, training, new culture) is often considered as more important than achieving a measurable outcome performance such as reduction in poverty (see Hugues and Traynor, 2000)

It is widely recognised that today public managers work under a context of change and uncertainty. Consequently, managers must have a capacity for readjustment and adaptation. This situation does not allow us to realistically assess organisational performance only on the basis of traditional performance measurements. There is a need for new measures that capture the capacity of managers to identify, respond and cope with unanticipated problems.

In any case, the 3Es and 3Ds approaches for evaluation are not incompatible alternatives. Using the traditional 3Es at least as 'tin-openers' - signals for further study - (Carter 1989) improves performance monitoring, which could be complemented with traditional program evaluation. But new measures or criteria of success are needed for assessing the management capacity for *Diagnosis* of new problems, for *Designing* new solutions and for *Developing* or getting support for their implementation. According to Metcalfe, the 3Ds approach should allow the assessment of:

Diagnosis capacity for the identification of new problems and the redefinition of current problems taking into account changes in the environment, the interests of stakeholders, and the need to build a common perception of the problem.

Design capacity for the formulation of new solutions (policies and institutions), including the incentives to solve new identified problems and the adaptation of organisational and inter-organisational structures and strategies.

Development capacity for the actual implementation of new solutions as a learning process, coping with resistance to change, redefining problems and solutions during implementation, and learning from experience. (Metcalfe, 1991):

The 3Ds model looks at innovative and adaptive performance and demands questioning the definition of objectives, the cause-effect relationship and the values underlining the public sector intervention. A parallel must be drawn here between, on the one hand, the 3Es vs. 3Ds

distinction and, on the other hand, the single vs. double loop organisational learning discussed in an earlier book in this series by Leeuw, Rist and Sonnichsen (1994) and the inter-organisational learning discussed by Metcalfe (1993).

The 3Ds approach is more relevant for assessing performance under uncertainty. Rather than examining how well performance criteria based on the 3Es have been met in cases where these criteria become obsolete or irrelevant, assessing performance from a 3Ds perspective looks at how well organisations and their managers confront unstructured contexts (see table above), i.e.:

- respond to multiple pressures and unexpected challenges (diversity);
- exhibit flexibility in the face of changing circumstances (instability);
- work co-operatively within networks of organisations when delivering new programs (interdependence); and
- adapt to an unpredictable environment (uncertainty).

In an unstructured context, we suggest that the central budget office should assess public managers' performance with regard to their capacity for adapting to new problems beyond the traditional concerns of cost reduction, productivity and quality improvement of current services. In such a context, maximising adaptability rather than maximising short term utility is a different and more relevant criteria of success.

A pragmatic decision rule for both budgeters and managers might be to build a balanced mix of both evaluation based on the 3Es and a set of evaluations based on the more relevant 3Ds. The choice of appropriate mix would depend on the degree of uncertainty under which managers work. It would, thus, become important to distinguish between various situations. On the one hand, and to the extent that agencies have straightforward and undisputed objectives and deal with well defined, homogeneous tasks and predictable outcomes, the 3Es approach and traditional evaluation will probably cover most of the management information needs. On the other hand, for organisations working in complex environments (with objectives subject to different views and interpretations, networks of organisations under policy change, central units

dealing with administrative reform) the need for a 3Ds evaluation approach would become higher.

Approaches to evaluation under uncertainty

The need for a more strategic leadership from central budget units leads to look for new types of evaluation approaches which can stimulate and reward innovation and experimentation in a complex public policy environment.

In budgeting, there is a need for explanatory material allowing to interpret and adapt the logical background of budget programme performance. In this respect, evaluation often appears to be the necessary tool that can provide the needed interpretation to try and sort out the various contributing factors to program performance

Measuring the 3Ds type of performance requires alternative evaluation approaches. In this sense, through decentralised evaluations focusing on participation, adaptability and collective strategy building, some weaknesses of the traditional centralised budget process can be corrected. The following points present an overview of some types of evaluation approaches which facilitate assessments under the 3Ds model and provide some hints on BO new roles. Other chapters in the book treat these evaluation approaches in more detail.

Pragmatic approaches: short timely information

The basic idea of providing usable knowledge corresponds quite adequately with the pragmatic approach defended by Thoenig and others (Thoenig, 2000). From this perspective, evaluation is considered to be more credible if it is adjusted to the reality of the decision-making process. Like corporate executives, public decision-makers give priority to practical or qualitative information obtained by speaking to individuals they trust. Therefore, evaluation should be limited in scope; focused on clearly defined problems; use language which policy makers and the society at large can understand; and obtain data that is available or can be obtained rapidly at a low cost.

In accordance with these characteristics, the basis for this type of pragmatic evaluation is that evaluators should not embark on making judgements and assessments (particularly when they must evaluate on-going reforms). As expressed by its advocates, evaluation should be decentralised and participatory, done concurrently by networks of professionals directly involved in the implementation of reforms and it should aim at describing a state of affairs rather than analysing it. As a consequence, evaluation will more likely provide performance proxies rather than assess external impacts. For BO, this approach may produce the kind of evaluative knowledge that can be easily shared and exchanged. It also allows BO to continue to play a monitoring role.

This evaluation approach is attractive for its simplicity. It assumes that knowledge will be quite explicit and relatively easy to handle. It also assumes that people have defined roles and that the flow of information is mainly bottom-up. However, innovative public policies often require understanding how it is possible to induce change through non-standardised processes. Often this will require a lot of dialog and the flow of information to be interactive.

Theory Based Evaluation: theoretical assumptions

Theory based evaluation was a response to the questions raised on the view of programs as black boxes (Chen, 1990; Weiss,1995). In a recent article of Birckmayer and Weiss (2000), we found a brief and precise description of this approach. According to them, the basic idea is to surface the assumptions on which the program is based in considerable details. The evaluation will have to review what activities are being conducted, what effect each particular activity is meant to have, what are the following steps in the programme, what the expected response is and so on, so as to find out whether the process matches the expected outcomes. The evaluation then follows each step in the sequence to see whether the expected steps actually materialise.

Following the same authors, an evaluation which investigates the theories underlying a programme will, thus, have to collect data at many points along the course of the programme so as to answer various questions, such as: does the program succeed in getting the attention of

target groups? Do the activities bring about a sense of trust? Does the program provide the planned activities? And this series of questions would end with a question on the outcome of the programme. Therefore, what TBE does is to track the stream of steps along the road so as to find out whether the theories on which the programme is based are realised in action .

The benefits that advocates of TBE claim for this approach (Birckmayer and Weiss (2000) are of three kinds:

1. TBE provides information about the mechanisms that intervene between programme activities and the achievement (or non achievement) of expected results. The evaluation can track each link in the chain of assumptions. As a consequence, the results of such an evaluation will show which chains of assumptions are well supported by data collected, which chains break down, and where in the chain they break down.

2. TBE is supposed to provide better knowledge about the mechanisms of change which will not only benefit the specific programme under study. The hope is that the knowledge will generalise to a wider array of change efforts. In this way, it enables the results from the evaluations to be fed into a wider knowledge of such programmes.

3. TBE also highlights the elements of programme activity that deserve attention in the evaluation, and it requires the operational staff to be explicit about their assumptions.

TBE is particularly adequate to know more about the actual effects of policies. Too much research and evaluation are focussed on implementation and participation, but not enough on the actual processes leading to outcomes of policies. The potential of TBE is there but so, too, is the risk of confusion in the collective sense of what is finally secure knowledge. Focussing on the theories of action gives evaluation a deeper sense that may be also conflicting with the time pressure and the need for quick answers of both managers and budgeters.

For BO, this approach may allow them to go beyond the straightforward evaluation of programs by agencies, producing the kind of evaluative knowledge that deals with the complex interactions which exist between various types of interventions and effects.

Participatory Evaluation: stakeholders constructions

Evaluation as a participatory process is based on the identification and involvement of stakeholders, the surfacing of claims, concerns and issues to facilitate consensus in negotiations. From this perspective, evaluation is seen as a political act. According to Stake (1975), the participatory evaluation is constructivist in the sense that perceptions on program activities and effectiveness are social constructions. It is also considered to be a responsive approach since parameters and boundaries for the evaluation are to be determined through an interactive and negotiated process involving stakeholders

One of the main arenas for participatory evaluation is in international cooperation. According to Jackson: "Participatory evaluation can be defined as the production of collective knowledge and cooperative action in which the stakeholders in development interventions participate substantively in identifying evaluating issues, designing evaluation, collecting and analysing data, and acting on the results of evaluation findings" This type of community-based participatory evaluations can play an important role as a leverage that might reorient public sector bureaucracy (Feinstein and Puetz, 2000).

Other similar well known approaches, such as Patton's utilisation-focused evaluation (1986) reduce the scope of participation by doing evaluations that provide information which decision makers view as necessary. This approach is, thus, based on the principle of usefulness and actual use of the evaluation. In this case, there is a need to identify the intended users of the evaluation, who are brought together and organised into an evaluation task force, to work with the evaluator and take part in the decision making of the evaluation. The evaluator together with the intended users decide the focus of the evaluation, review the options and select the approaches. The

evaluator helps the intended users to decide whether the evaluation is worth doing. If they decide to proceed, next stages embrace discussions about methods, measurement and design. Further on, data are collected and analysed and the intended users actively participate in interpreting findings, forming judgements and making recommendations. In the light of findings, it is then decided whether the evaluation results are to be disseminated.

This type of evaluation approach does not foresee the participation of all stakeholders since it is considered that participants have to be enthusiastic, committed, competent and interested. The idea is that the primary participants are lower down in the hierarchy and that participation is bottom-up. Therefore, not all the participants are motivated enough nor have the ability to support their involvement. Nevertheless, Patton considers that ability is not an impediment as such and that the project might have a low level of participation.

Participation is widely believed to be positive. Participation gives explicit consideration to social, political and cultural factors since it is based on a need to include, discover and understand. From a philosophical point of view, participation is also viewed as necessary given the interdependence to create knowledge. That is, each actor depends on others to provide information and to understand its own position.

For BO, this approach to evaluation will be difficult given the uncertainty associated, not only with the policies but also with the evaluation process itself. According to Gregory, there may be an imbalance of knowledge between stakeholders; fundamental conflicts may arise which are not easily resolved and; the free flow of information between stakeholders may be an ideal that is rarely achieved. Still BO can play a role in the diffusion of those experiences that appear to be successful from the perspective of the stakeholders.

Building trust for evaluation knowledge utilisation

Before we get carried away by evaluation being integrated with decision-making (by being more pragmatic) or more theoretical (in looking at theories of action) or more diverse and

interdependent (because of its participatory nature), we need to think about the deficit of trust between policy actors.

As it has been pointed out above, a long bureaucratic and incremental tradition has created deeply rooted culture of control and budgeting based on inspection of spending behaviour. As a consequence, in order to reinforce the shift from traditional to strategic budgeting, there is a need to break the vicious circle originated in the inherent criticism from controllers, and the defensive attitudes from public managers. Only if relations among them become constructive instead of competitive, sharing evaluation knowledge will be feasible. Public managers should perceive reports and constructive criticism as normal step towards organisational learning (OECD, 1996).

The renewed focus on social capital provides a framework for evaluation knowledge management. For the majority of writers (Bourdieu 1986, Coleman, 1988, Putnam, 2000, OCDE 2001,), social capital is defined in terms of networks, norms and trust and the way these allow agents and institutions to be more effective. Reciprocity and mutual support norms are basic for the construction of social capital. Historical personal experiences are as influential as the set of formal rules in determining the potential for interaction and validation of knowledge.

Building trust in the budget procedure depends on long-lasting and continuous relations of conditional and reciprocal nature. Co-operation should be oriented towards a common objective: the flow and use of evaluative knowledge. Trust development should facilitate the perception and awareness of interdependence in the network of managers, budgeters and evaluators. These would reduce strategic uncertainty, risk levels and the possibility of undisciplined budget behaviour.

Institutional structures for evaluation knowledge management require the existence of well functioning forums and organisms that facilitate the connection and integration among budgeters, managers and evaluators. This organisational mechanisms should be coherent with the context. For instance, in a decentralised and uncertain context, horizontal interaction could be more effective than hierarchical channels.

Building trust is a slow and a costly process whereas its destruction can be quick. This long process of generating mutual trust among of all the actors involved could be led by the department of the budget, or any other unit but should count on the support from other departments. Some concrete initiatives for developing social capital among evaluators, managers and budgeters are the following:

- Strengthening evaluation and budgeting units
- Job mobility, resource sharing and training programs
- Joined formal and informal activities for evaluators, budgeters and managers.
- Explicit declarations of willingness to encourage innovation, to risk failure and to learn from experience
- Accepting a diversity of evaluation models and approaches
- Forums for policy discussion

Social capital is essential for organisational learning and innovation in the public sector. A substantial increase of evaluation knowledge could be ineffective and even an obstacle disorienting budgeters if it does not go along with a corresponding development of trust and social capital in the budget process. Transforming evaluation information into knowledge ready for decision making and resource allocation demands a balance between information produced and existing collective trust in an organisation. The availability of data, figures, and information on results is insufficient if it is not accompanied by the necessary capacity (technical and social) for collectively selecting and transforming this information into knowledge.

Positive oriented accountability

Another important aspect of the new role for the budget office is related to spending control and accountability systems. The traditional aim of financial control is to fight fraud. As Metcalfe and Richards suggest, this aim should be complemented with a more positive perspective of encouraging the desired overall budget performance (Metcalfe and Richards, 1990). Because of

recent spending decentralisation trends, initiatives to improve financial accountability systems need to be reinforced to ensure that delegated spending authority is being applied properly and results have reached anticipated standards. This reinforcement does not mean just increasing the capacity to identify culpability or to enforce punishment.

To be successful in the improving of accountability, it is necessary to change the commonly held but false assumption that accountability and spending performance are inherently in conflict. Accountability is usually reduced to issuing and applying legal and financial regulations to constrain the efficient performance of public servants. Although difficult to achieve, a well-designed accountability system may encourage good spending results without damaging public service ethics. There is not necessarily a trade-off between accountability and performance. What is needed is to overcome the widespread negative view of accountability. This applies to both those in charge of running budget accountability institutions and budget holders. The objective of financial control should not only be to prevent irregularities in the use of public resources but also to motivate desired spending behaviour. And vice versa, performance management control may be useful to find out and correct irregularities.

A positive perception of accountability might encourage desired budget decisions rather than just prevent undesired ones. In other words, the two faces of the budgetary accountability coin should be taken into account. On the one hand, there is the constraining aspect of checking poor spending performance or the misuse of power (accountability institutions reacting negatively to correct unacceptable budget non-discipline). On the other hand, there is the guiding approach which requires the provision of a proactive and continuous feedback by setting values and norms from which policy and standards are derived.

Budget reforms since the 80s prescribe budget targets and performance standards. However, solving the deficiencies of current accountability institutions is not just a question of finding ideal criteria (performance measures) for success. The effectiveness of accountability systems has been diluted to mere answerability, making it only necessary to justify and explain what has been done and to respond to questions raised by Parliament (Metcalfe and Richards ,1990).

One of the difficulties lies in recognising and capturing the multiple and conflicting pressures coming from different groups (managers, professionals, clients, peers, constituencies, etc.) claiming a need for modifications to spending targets. Ideally, priorities need to be established according to the relevance of such groups – as sources of legitimacy – to construct a framework of performance reference that synthesises all relevant criteria of success. Appropriate accountability regimes should be taken into account and be coherent with the existing sources of influence.

Accountability and budget performance assessment in the public sector should also be oriented towards adapting spending values and beliefs rather than just applying rational criteria (such as the 3 Es) and enforcing compliance. Negative accountability should be compensated with positively oriented assessments and organisational-program learning. Inspection and rigid control on undesired spending do not guarantee improvement in spending performance. On the contrary, at a certain point accountability can actually begin to encourage dysfunctional behaviour ('creative' accounting, biased or fictitious budgeting, concealing information, non-productive internal competition, etc.). In contrast, positive accountability facilitates the internalisation of sound performance values and provides guidance for sound spending behaviour.

In short, another role of the evaluation for strategic budgeting should be to facilitate a balanced functioning of the two faces of spending accountability systems. This means avoiding monopolising accountability efforts on inspection and negative performance in which the only concern of the central budget unit is to keep the control and stability of the budget system. The traditional accountability system only functions when reacting to non-disciplined budget behaviour. A new accountability approach which is more proactive and continuous needs to be developed if capacity to change is considered a priority. The central budget unit and financial management units in spending departments should encourage a more strategic budgeting by explicitly or tacitly rewarding adaptive behaviour, i.e. taking risks and experimenting. In a context of instability, a positive perception of accountability as guidance rather than inspection is more appropriate since an overemphasis on inspection of non-desired budget behaviour may

mean blockage and incapacity to adapt. In this respect, a continuous guiding role played by the central budget offices facilitates adaptation (Mayne and Zapico, 1997)

III.   CONCLUSIONS

For some years now, in most OECD countries, it has been claimed that budgeting decisions should be supported by performance information and that results management was need to promote a change in the culture of government. Despite some progresses in performance measurement and assessment, budgeters and managers have rather focussed on technical aspects of designing measurement indicators. Depending on the kind of services and programmes, indicators have been valued or simply considered a bureaucratic response to the demand of evaluative knowledge. Single evaluation studies have also been considered, at times useful to clarify issues, but more often a formal response to an obligatory requirement.

The need for change and innovation in performance measurement and evaluation is particularly intense if we take into account the increasing tendency towards a decentralised agency model and the existence of a complex public policy environment. In a decentralised system, the traditional type of budgeting control and "negative" accountability needs to be put into doubt. Decentralisation calls for a more flexible type of spending control based on the understanding of problems and contextual factors. This is even more the case in those instances of public service that operate in an unstructured context where there is uncertainty, instability, diversity and interdependence.

In this chapter, strategic leadership from BO has been defined as contributing to develop the kinds of evaluative knowledge that are most relevant to budgeters, managers and stakeholders. A pragmatic decision rule for BO can be whether evaluation should be based on the traditional 3Es or on the ideas behind Metcalf's 3 Ds. The basic distinction is that the more traditional 3 Es approach assumes that predefined and stable objectives are available which presupposes a well-known cause-effect relationship between organisation activities and outcomes. However, the capacity of managers to be Economic, Efficient and Effective is insufficient when goals are changing or when outcomes cannot be readily anticipated.

In a more complex environment where public managers work in a context of change and uncertainty, the 3Ds approach is better equipped to capture the innovating capacity of public managers. That is the capacity to carry out the Diagnosis of new problems, the Design of new solutions and to Develop their implementation.

When turning to evaluation theories and approaches in a context of uncertainty, we conclude that there is not one single formula or model but a variety of avenues and sources of information. The key issue is that the information is available and that it is conducive to learning since learning through evaluation is building social capital among budgeters, managers and stakeholders. Learning may require practical information, limited in scope, focussed on clearly defined aspects of a problem, rapidly available and at a low cost. It may also be the compensation for efforts to disentangle the underlying -"theoretical"- assumptions on which a program is based. Learning can also come as the sudden understanding of processes through information produced by stakeholders as a response to their questions and concerns.

But the production of information is not enough. For actors to play a role or act in a knowledgeable manner, they need to change their interaction patterns. The BO has a role to develop in the creation of structures, rules and incentives that may facilitate a better balance between competition for resources and cooperation and mutual understanding. In particular the change in the nature of the relations among budgeters, managers, evaluators and other stakeholders would be encouraged by a change in the accountability systems, from the traditional inspection role that reacts to non-disciplined budget behaviour, to a proactive role that explicitly encourages and supports strategic spending behaviour.

Learning and social trust will benefit from exchange and participation. It is crucial to realise that other actors can see things in different ways. Very often, what seems non-rational or inadmissible at a first glance will later find its place and show its potential.

References

- Argyris, C. (1982). *Reasoning, Learning, and Action*. San Francisco: Jossey-Bass Publishers.

- Birckmayer, Johanna D. and Carol H. Weiss (2000) "Theory-Based Evaluation in Practice: What do we learn?" *Evaluation Review,* Vol. 24 (4): 407-431.

- Bourdieu, P. (1986) "The Forms of Capital" in J. E. Richardson, (ed.) *A Handbook of Theory and Research for the Sociology of Education*, New York: Greenwood.

- Canadian Comprehensive Auditing Foundation (1987). *Effectiveness Reporting and Auditing in the Public Sector*. Ottawa

- Carter, N. (1989). "Performance Indicators: Backseat Driving or Hands off Control ?" *Policy and Politics*. XVII(2): 131-38.

- Carter, N. 1991. " Learning to Measure Performance: The use of indicators in Organizations" in Public Administration 69.

- Chen, H. (1990) *Theory-Driven Evaluation*, Newbury Park, CA: Sage

- Coleman, J. (1988) "Social Capital in the Creation of Human Capital", *American Journal of Sociology*, Vol. 94 Supplement 95-120.

- Derlien, H.U. 1990. "Genesis and Structure of Evaluations in Comparative Perspective," in R. Rist (ed.), *Program Evaluation and Management of Government*. New Brunswick: Transaction Publisher

- Fernandez Ballesteros, R (ed), 1995 *Evaluación de programas. Una guía práctica en ámbitos sociales, educativos y de salud*. Madrid. Editorial Síntesis.

- Feinstein, O. and Puetz, D. 2000. "Synthesis and Overview" in O. Feinstein and R. Picciotto "Evaluation and Poverty Reduction", The World Bank, Washington DC..

- European Group of Public Administration (1996) "Strategic Spending Management A new role for the Budget Centre", Budapest meeting 1996.

- Gregory, Amanda 2000. "Problematizing Participation", *Evaluation*, Vol 6 (2): 179-199

- Gunther, R. 1980. *Public Policy in a No-Party State Spanish Planning and Budgeting in the Twilight of the Francoist Era*, p. 51. Berkeley, CA: University of California Press.

- Hugues, M. and T. Traynor (2000) "Reconciling Process and Outcomes in Evaluating Community Initiatives" in *Evaluation*, Vol 6 (1): 37-49

- Jackson, E. (2000) "The Front-End Costs and Downstream Benefits of Participatory Evaluation", in O. Feinstein and R. Picciotto "Evaluation and Poverty Reduction", The World Bank, Washington DC.

- Leeuw, F. L., R. C. Rist, et al. (1994). *Can Governments Learn? Comparative Perspectives on Evaluation & Organizational Learning*. New Brunswick, N. J.: Transaction Publishers.

- Mayne, J. and Zapico E. 1997, *Performance Monitoring in the Public Sector*, London: Transaction Publishers, Chapters 1, 4, 2 and 8.

- Mayne, J. (1994). "Utilizing Evaluation in Organizations: The Balancing Act". In F. L. Leeuw, R. C. Rist and R. C. Sonnichsen (Eds.) *Can Governments Learn? Comparative Perspectives on Evaluation & Organizational Learning*. New Brunswick, N. J.: Transaction Publishers.

- MEANS, 1993 ???? "Les impacts sur l'emploi dans le secteur coopératif catalan: Project dánalyse factorielle" in *MEANS 1995-1: Analysis of ex-post evaluations of the Community support frameworks of objective 1 regions for the period 1989-93. Methodological lessons across seven Member States*. Lyon: Means program.

- MEANS, (1995-2): For improved evaluations of Community Programmes. Results of the 1995 MEANS program. Homogenizar con la anterior

- Metcalfe, L and S. Richards (1990). *Improving Public Management*. 2nd edition. London:Sage, pags. 210, 96, 40, 44.

- Metcalfe, L. (1991). "Public Management: From Imitation to Innovation". Paper presented at the Annual Conference of the International Association of Schools and Institutes of Administration (IASIA), Kota Kinabalu, Sabah, Malaysia, 28 July - 2 August. 9-10.

- Metcalfe, L.(1993) "Public Management: From Imititation to Innovation" in J. Kooiman. *Modern Governance*. Londres: Sage Publications.

- Monnier, E. (1995) *Evaluación de la acción de los Poderes Públicos*. Madrid, IEF.

- Murray, R. (1992). "Modernization of Public Sector in Sweden". Paper presented in the conference *Public Management Modernization in the 90s*, INAP, Madrid: 2-4.

- OECD (1993) *Performance Measurement*.. PUMA (3) 93. Paris

- OECD (1993) *Public Management Developments, Survey 1993*, Pages 5-11.

- OECD (1994) *Service Quality Initiatives in OCED Member Countries*. Background Paper for a Symposium held 7-8 November, 1994. PUMA (94) 13.

- OECD, (1996). Performance Auditing and the Modernization of Government, PUMA, Paris.

- OECD (2001) "The Well-Being of Nations: The role of Human and Social Capital", Paris.

- Office of Management and Budget (1995). *Making Government Work: Budget of the United States Government, Fiscal Year 1996*. Washington DC.

- Patton, M.Q. (1986) *Utilization-Focussed Evaluation*, London, Sage.

- Perrin, Burt, (1998)"Effective Use and Misuse of Performance Measurement" *American Journal of Evaluation*, Vol. 19 (3): 367-369.

- Putnam, R. (2000) *Bowling Alone: The Colapse and revival of American Community*. Simon Schuster, New York

- Restrepo, P., Arribas, M., y Torrego, J.M., (1990) *Evaluación del trabajo y resultados*, Madrid, Consejería de Integración social de la CAM.

- Rist, R (ed), (1996) *Program Evaluation and Management of Government*. New Brunswick: Transaction Publishers

- Savoie, D. J. (1994). *Globalization and Governance (Summary).* Research Paper No. 12. Canadian Center for Management Development. Ottawa.

- Selznick, P. (1957). *Leadership in Administration*. Illinois: Row, Peterson and Compan

- Shandal, R. (1999) "Performance-Monitoring Systems: A Basis for Decisions?" in J. Mayne J. and E. Zapico-Goñi *Monitoring Performance in the Public Sector*. London: Transaction Publishers

- Schick, A (1990) Budget for Results: Recent Developments in Five Industrialized Countries". *Public Administration Review*. Vol (?????)

- Smith, D. (1978). *Value Biases in Performance Assessment*. Paper presented at the 1978 Annual Meeting of the American Political Science Association in New York

- Task Force on Management Improvement (1992). *The Australian Public Service Reformed - An Evaluation of a Decade of Public sector Reform*. Canberra: Australian Government Publishing service.

- Thoenig, Jean-Claude, (2000) "Evaluation as Usable Knowledge for Public Management Reforms", *Evaluation*, Vol. 6 (2) 217-229

- Weiss, Carol, H. (1995) "Nothing as Practical as Good Theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families". In J. P. Connell, A.C. Kubich, L. Schorr, and C. Weiss, *New Approaches to Evaluating Community Initiatives: concepts, methods and contexts*. Washington, DC: Aspen Institute.

- Wildavsky, A. (1993). "¿Qué necesita saber el manger público?' Una formación para la gestión pública". Ekonomiaz, 26(2) cuatrimestre. Departamento de Economía y Hacienda

- Wildavsky, A. (1975) *Budgeting: A Comparative Theory of Budgeting*, p. 4. Boston: Little Brown

- Wollman, H. (1984) "Policy Research in Spain," in *Development and Present State of Public Policy Research: Country Studies in Comparative Perspective*, study discussed at a workshop at the Wissenschaftszentrum, Berlin, December, 1984, pg. 709.

- Zapico, E. (1989), La Modernización Simbólica del Presupuesto Público, HAEE/IVAP, pg.145