Program Performance Monitoring and Evaluation
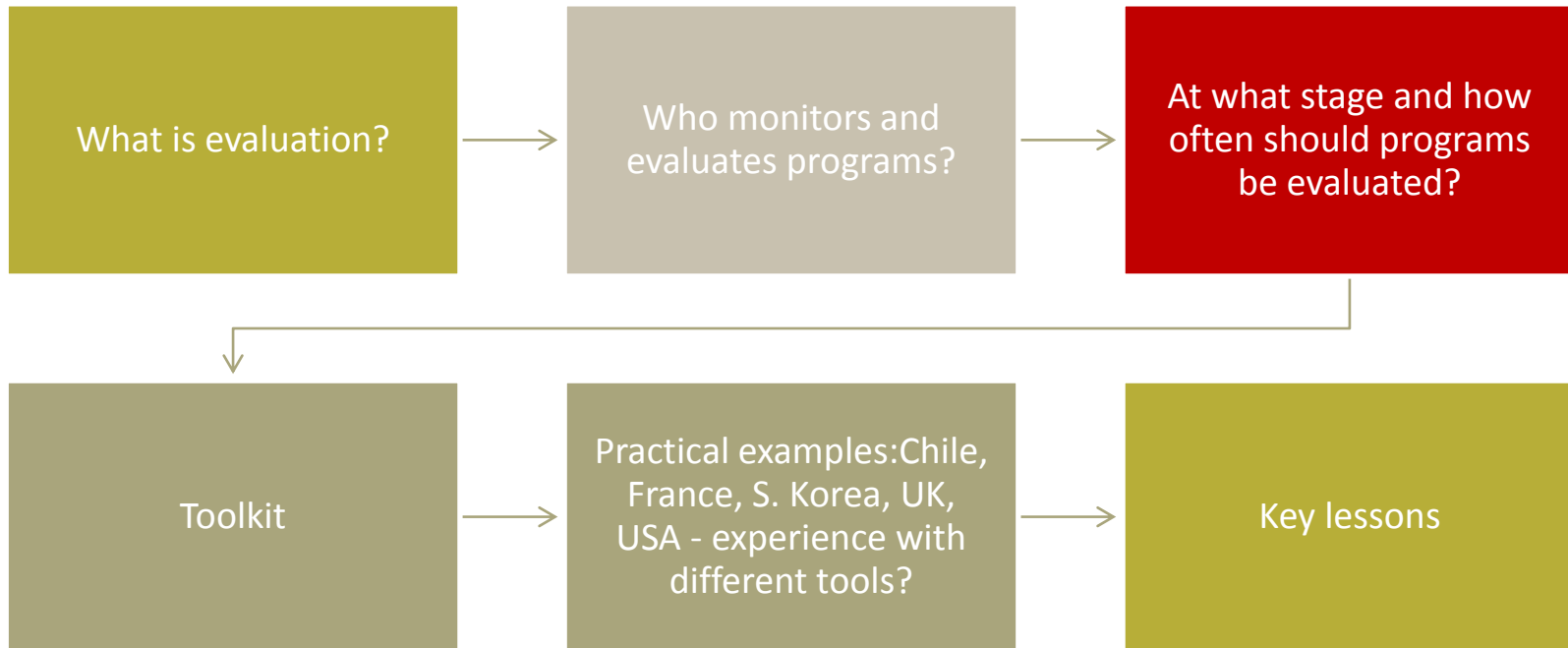
Ivor Beazley

BCoP, Tirana, February 2013

Theory and practice

- Program Budgeting links financial resources to results and performance.
- For the system to work there needs to be a process for evaluating results and performance
- Did the programs fulfill their objectives. If not, why not and what should be done about it?
- Results of an effective the evaluation process should be:
  - better resource allocation,
  - better program design and management
  - better results
- In practice evaluation is often the weak link in performance budgeting

What we will cover

What is evaluation?	Who monitors and evaluates programs?	At what stage and how often should programs be evaluated?

Toolkit	Practical examples:Chile, France, S. Korea, UK, USA - experience with different tools?	Key lessons

What is evaluation?

Definition: *The systematic and objective assessment of an on-going or completed project, program or policy, its design, implementation and/or results. Evaluation aims include the relevance and fulfillment of objectives, efficiency, effectiveness, impact and sustainability.*

Formative evaluations strengthen or improve the object being evaluated

- *Needs assessment* determines who needs the program, how great the need is, and what might work to meet the need
- *Evaluability assessment* determines whether an evaluation is feasible and how stakeholders can help shape its usefulness
- *Structured conceptualization* helps stakeholders define the program or technology, the target population, and the possible outcomes
- *Implementation evaluation* monitors the fidelity of the program or technology delivery
- *Process evaluation* investigates the process of delivering the program or technology, including alternative delivery procedures

Summative evaluations examine the effects or outcomes of some object

- *Outcome evaluations* investigate whether the program or technology caused demonstrable effects on specifically defined target outcomes
- *Impact evaluation* is broader and assesses the overall or net effects -- intended or unintended -- of the program or technology as a whole
- *Cost-effectiveness and cost-benefit analysis* address questions of efficiency by standardizing outcomes in terms of their costs
- *Secondary analysis* reexamines existing data to address new questions or use methods not previously employed
- *Meta-analysis* integrates the outcome estimates from multiple studies to arrive at an overall or summary judgment

Objectives of evaluation

Strategic

- Are we doing the right things?
- Program rationale/justification
- Are the objectives clear and realistic?
- Are the outputs appropriately defined and measured?
- Has the linkage between outputs and outcomes been adequately established?

Operational

- Are we doing things right?
- Effectiveness - are the benefits reaching the intended beneficiaries?
- Does the program offer value for money
- Is the program adequately funded to meet its objectives
- Are controls effective
- Is the data and reporting reliable?

Learning

- Are there better ways of doing it?
- Best practices
- Lessons learned
- Alternative designs and delivery mechanisms

Who monitors and evaluates and with what objectives?

KEY ACTORS	Their interest in in-year monitoring	Their interest in ex post evaluation
Presidency or Cabinet	Monitoring important/sensitive programs. Use of monitoring information is particularly strong where there are "performance agreements" between the president and line ministers or between ministers and ministries and other agencies.	Program performance with respect to achievement of key policy objectives (non financial performance) and outcomes Evaluate Ministers' performance in program delivery (PSA) and hold them accountable. Identify and unblock obstacles to program delivery.
Planning Agency	Track in-year performance of public investment programs.	Evaluate progress against the National Development Plan objectives Assess program and policy effectiveness.
Finance Ministry	Monitor in year program performance including risks affecting fiscal position. Some finance ministries monitor non financial performance targets and performance agreements. Can also be used as part of performance management and review key staff.	Performance information can be included with the annual financial statements submitted to the legislature and published after the end of fiscal year – with implicit or explicit evaluation of program effectiveness
Line Ministries	Monitoring data (outputs and outcomes) can indicate performance of ministry and agencies and can enable policy adjustments (e.g. shift in priorities, program re-design, re-allocation of resources) with a sector. Also used in internal management of the line ministry, and in the context of supplier-provider arrangements or output based disbursements within the public sector or with other entities, including through use of league tables.	Annual reports by ministries can incorporate performance as well as financial information. Program evaluation can both draw on and contribute to performance information. Line ministries can be required to improve the quality of data and of measurement on the basis of the findings and recommendations of audit reports

Continued

| KEY ACTORS	Their interest in in-year monitoring	Their interest in ex post evaluation
Legislature		There is some reporting of performance in budget execution to legislatures in many OECD countries, allowing some degree of program evaluation by the legislature. Ministers make plans for their ministries partly in terms of outputs, and can be held accountable for them in the Legislature. It is important to note, however, that in most OECD settings, failure to deliver output targets generally forms the basis for a discussion, not an automatic sanction.
External Audit	N/A	Performance auditing informs the legislature and other stakeholders about program performance and management. Can assess the quality of program management, controls and reliability of performance information Assessment of value for money
The Public	Organized groupings can use aggregate performance measures to push for better execution in policy areas that concern them. Individuals can use information concerning compliance with service standards in improving budget execution/seeking redress in relation to particular services.	Performance information can be used by the public to assess government policy and program delivery against promises made, formal commitments in law and accountability documents. League tables, benchmarking, citizens' charters etc. can enable individuals to assess performance of service providers.

Program review cycle

Annual
- Ex ante scrutiny – part of annual budget process
- Regular (annual) monitoring and reporting of program expenditure and results

Periodic - Comprehensive
- Comprehensive review across programs so that they can be compared and strategic choices made –e.g. UK Comprehensive Spending Review, WB Public Expenditure Reviews (once in a parliament)

Rolling - partial in –depth coverage
- Rolling program of in depth monitoring – focused on program design and effectiveness – e.g. PART in USA(OMB)
- Independent performance audits,
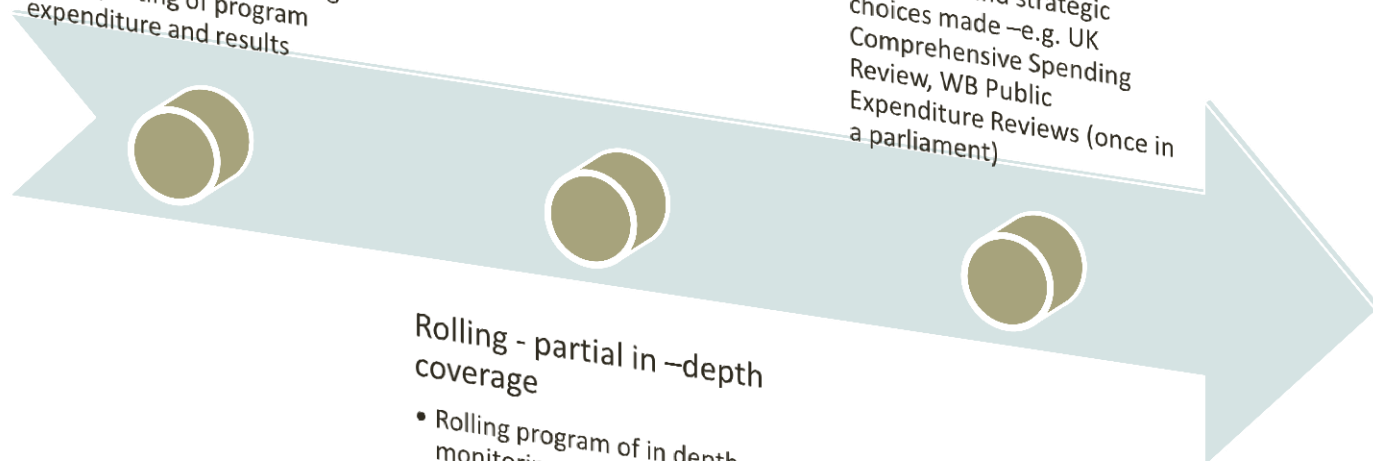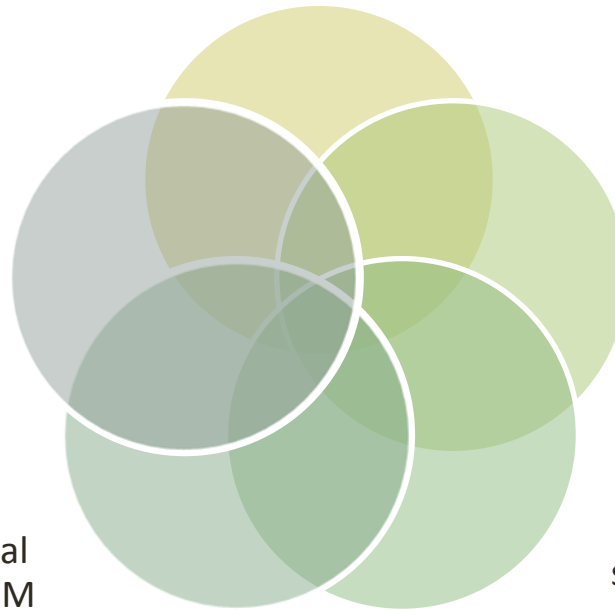
Performance Evaluation Toolkit

Ex ante scrutiny – mid term and annual budget processes

Centre of Government oversight unit (MoF, President's Office) – e.g. PM's delivery unit.

Ministry \self-evaluation e.g PART

Independent external performance and VFM audits - (NAO, CoA)

Cross government spending reviews (UK Comprehensive Spending Reviews)

Case Study: USA -Program Assessment Rating Tool (PART)

- During the last 20 years, Congress and the Executive Branch have sought to instill a greater focus on results

  - The Government Performance & Results Act (GPRA) of 1993 was the centerpiece of the statutory management framework of the 1990s
    - Key purpose: to create a closer and clearer connection between resources and results
    - GPRA requires all of an agency's budget program activities to be covered in its performance plan.
    - Prior initiatives failed in part because they weren't relevant to congressional and executive budget decision making.

Government programs are very varied, requiring different approaches to performance evaluation

- **Competitive Grants -** Distribute funds via competitive process.
- **Block/Formula Grants -** Distribute funds by formula.
- **Regulatory-Based -** Achieve goals through rulemaking.
- **Capital Assets & Service Acquisition -** Achieve goals through the development & acquisition of capital assets or the purchase of services from a commercial source.
- **Credit -** Provide support via loans, loan guarantees, direct credit.
- **Service provision –** Service provided mainly by government employees.
- **R&D -** Develop knowledge or apply it toward the creation of systems, devices, methods, materials or technologies.

GAO Evaluation of PART

PART process aided OMB's oversight of agencies, focused agencies' efforts to improve program management, and created or enhanced an evaluation culture within agencies.

PART was a labor-intensive process at OMB and agencies.

Most PART recommendations did not result in short term performance improvements and provided wide latitude for agency responses

- PART increased expectations and invited reaction, but it also presents institutional challenges for the Executive Branch and Congress:
  - -- Whose framework, interests & perspectives should drive the process
  - -- How to get consensus or buy-in by stakeholders, especially Congress

Because of limited agreement between OMB and Congress about PART, congress did not use PART Information

Case Study; UK PM's Delivery Unit

Why?

- Explosive growth in KPIs and performance indicators
- Lack of responsiveness to key policy priorities
- Set up
- Unit based at the PMO, direct support of the PM, staffed by 40-50 Civil Servants but headed by a high profile outsider
- Remit to deliver on 30 key Public Service Agreements (in Health, Education, Criminal Justice and Transport). Ministers held personally accountable for PSAs.
- Clear and ambitious targets for key services (embodied in 30 Public Service Agreements (PSAs)).
- Delivery Unit offered expertise and methodology. Worked with departments to agree 'trajectories' to meet PSAs.
- Defined appropriate indicators by which to judge success

Delivery units: Pros and Cons

Attractive model:

- The model delivered results and helped to unblock
- Outcome level targets drive performance improvements in core public services
- Delivery unit helps setting trajectories and approaches and holds institutions accountable for delivery, short cut change processes

But:

- Results often at the cost of rising expenditure
- Sustainability (and fatigue) is an issue, especially as gestation period is likely to be long in less advanced public management systems
- Needs strong and consistent leadership drive
- Black box (needs high internal capacity to manage)
- Competing systems of target setting (e.g. Malaysia and Indonesia retained their conventional planning systems)

14

NAO review/audit of program performance

- SAIs are well placed to contribute to performance improvement, but this often requires big changes in approach and shift of resources
- Strong SAIs are leading proponents of improvement – (e.g. Canadian OAG and US GAO) drawing attention to deficiencies in transparency, accountability, control and reporting systems and recommending improvements
- Some see commenting on the design of programs as compromising their independence to review/criticise,
- Lack of expertise in technical areas and budget are constraints – coverage limited
- Performance auditing is now an established feature of NAO work in OECD countries
- SAI coverage may include evaluation of Performance Budgeting across Government (e.g. Australia NAO assessment of the use of performance information, UK review of data supporting PSAs, GAO review of PART program)

SAI – selected country example examples

- **France – Court of Accounts** – ensures that results reported by each ministry/agency annual performance report are verifiable and CoA provides guidance

- **Australian NAO** Comments generally on the quality of performance information and reviews the impact of its performance audit reports, in particular linkage with government budgetary decisions

- **Swedish NAO** reviews (but does not formally attest) performance indicators and indicates cases where it considers the information is not "true and fair". Plays a consultative role, assisting agencies to improve their performance information

- **Netherlands NAO** expresses a formal audit opinion on the quality of non financial information in the government annual report . It has commented generally on inadequate quality of performance indicators

South Korea

Korea took a gradual approach in developing its performance monitoring and evaluation system. Korea has three layers of monitoring and evaluation system: monitoring, review and in-depth evaluation:

In 2000, it started a pilot project with selected departments in some line ministries to prepare **annual performance plans and reports**. In 2003, it was extended to all ministries.

In 2005, the **program review process** was introduce. This was an development of f the PART (program assessment rating tool) in the US to provide more systematic and comprehensive information on the performance of spending programs. It is called the "Self-Assessment of Budgetary Programs (SABP)"

In 2006, an **in-depth evaluation process** was introduced which is conducted on selected programs each year.

Among these systems, evaluation results from the program review process is more systematically and actively used in the budget allocation process.

Program Review in Korea

Program Review Criteria

Design and Planning (30)	• Program purpose • Rationale for government spending • Duplication with other programs • Efficiency of program design • Relevance of performance objectives and indicators • Relevance of performance targets
Management (20)	• Monitoring efforts • Obstacles of program implementation • Implementation as planned • Efficiency improvement or budget saving
Results and accountability (50)	• Independent program evaluation • Results • Utilization of evaluation results

The Program Review tries to make a link between evaluation results and budget allocation. The annual guidelines of budget formulation issued to line ministries by the central budget authority states that line ministries should consider at least a 10% budget cut to the programs rated as "ineffective" by the SABP.

Australia

Australia uses many forms of program evaluation depending on the nature of the issue under discussion:
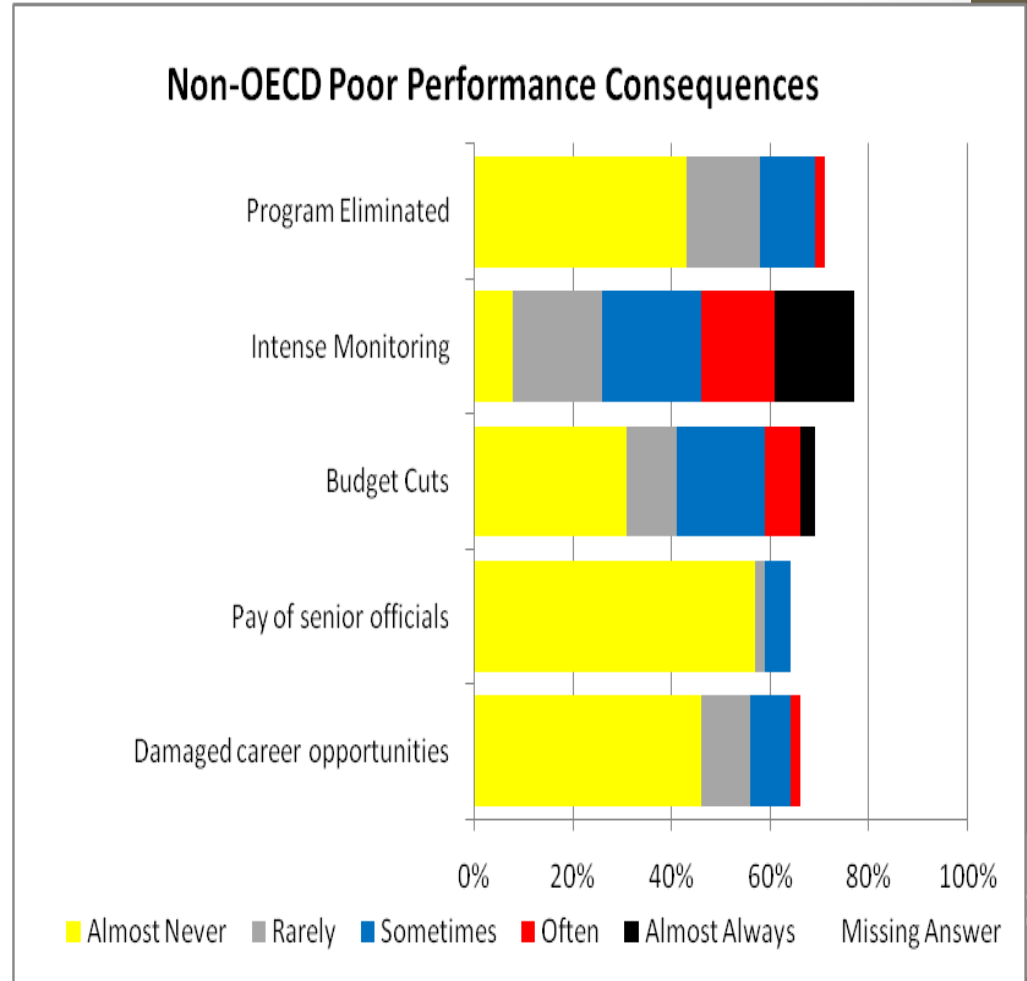
- ✓ **Price reviews**: focus on the economic efficiency of an agency's outputs and the price of those activities

- ✓ **High level specialist evaluations:** are often conducted by independent bodies (consultants, academics etc) to assess the impact of programs/outputs on outcomes and their contribution to the government identified goals

- ✓ **Cyclical evaluations:** (roughly every 3 years) test whether outputs and programs are still relevant to the purpose for which they were originally designed.

- ✓ **Internal agency evaluations:** are often generated to test whether the monitoring and KPI setting for a particular program are still in place and properly aligned to service delivery activities and outputs.

There has not been an attempt to establish a single online monitoring IT system that capture at a central level all data relating to all agencies.

Consequences of poor performance

- **Monitoring, Problem Solving and Support**

- Management Interventions

- Publication & League Tables

- Budget Incentives and Sanctions

- Personnel Incentives and Sanctions

More

Less

Credibility ? Effectiveness ?

Non-OECD Poor Performance Consequences

Program Eliminated	
Intense Monitoring	
Budget Cuts	
Pay of senior officials	
Damaged career opportunities	

0% 20% 40% 60% 80% 100%

Almost Never Rarely Sometimes Often Almost Always Missing Answer

20

Conclusions

- Program evaluation is and essential but often weak link in performance budgeting
- Monitoring and evaluation should be both ex ante and ex post.
- There is a wide range of tools and stakeholder - no single tool is effective on its own
- Self evaluation by program managers is critical to make timely course corrections
- Central Government scrutiny can be effective in maintaining focus on policy objectives and ensuring responsiveness
- Periodic review across all programs advisable
- External audit may need re-tooling to be effective in program review
- In depth evaluation is needed to determine reasons for poor program performance
- Decisions must generally be made with imperfect information - avoid "paralysis by analysis"
- Resources are limited – in depth reviews must be periodic
- Continuing problem to ensure that evaluation leads to action