OECD's work on RTD Evaluation

ESF Member Organization Forum on Evaluation

Rome, 31st March – 1st April 2008
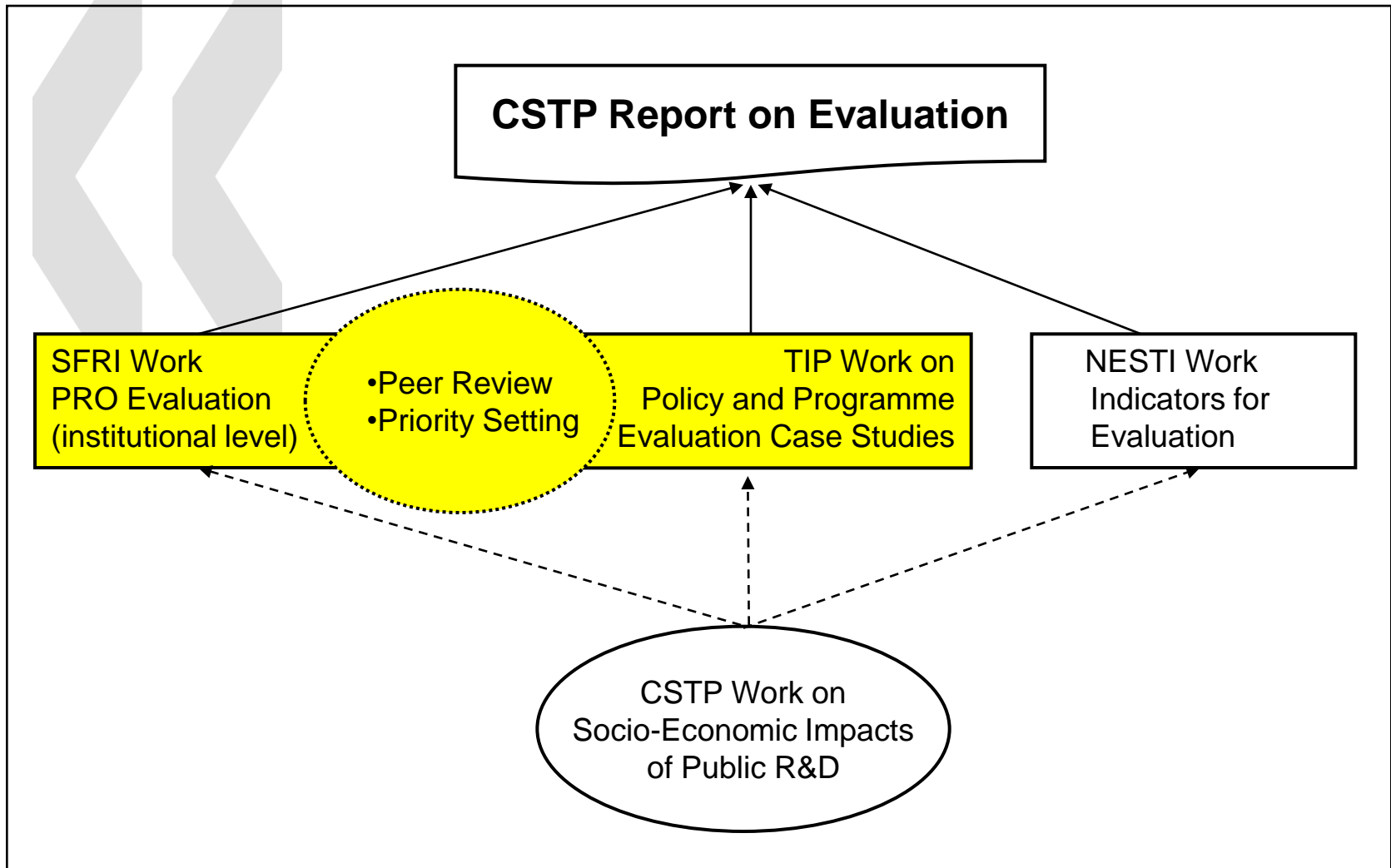
Dong Hoon Oh

dong-hoon.oh@oecd.org

OECD 《 1 OCDE

Introduction

● This presentation is expected to:

1. Introduce the CSTP's project on evaluation

2. Provide brief results of the TIP/SFRI joint workshop on **Peer Review** and **Priority Setting**

3. Introduce a **framework** and **guidelines** for case studies

4. Outline the **status** and **next steps** of the project

Framework of CSTP Evaluation Project

CSTP Report on Evaluation

SFRI Work
PRO Evaluation
(institutional level)

- Peer Review
- Priority Setting

TIP Work on
Policy and Programme
Evaluation Case Studies

NESTI Work
Indicators for
Evaluation

CSTP Work on
Socio-Economic Impacts
of Public R&D

OECD 3 OCDE

M1: Architecture of Evaluation Systems

- To examine the systems or the "architecture" of evaluation systems in different OECD countries.

- A **case study** approach can be an effective means
  - The frameworks will serve to gather country-specific information on the other modules of the project.

- By country case studies of volunteer countries:
  - Method
    - *Case studies* & *workshops* mainly in 2007.
  - Deliverables
    - *Compendium of good practices for evaluating research and innovation policies* to be published in 2008.

M2: Peer Review (Expert Review)

- "Peer review" is one of the most common methods.
  - PR is quick, clear, practical, and useful for mutual learning.
  - PR is under pressure and losing confidence among users.

- How to combine PR with objective indicators?
  - Can or should the peer review system be replaced?
  - Is evaluation possible without peer review panels?
  - What kind of PR is fit for the evaluation of an innovation policy?

- The work could be done parallel to the case studies.
  - Method: Small workshop with practitioners
  - Deliverables: 1) Best Practice, 2) Recommendations

M3: Priority Setting and the Use of Evaluation

- PS is an important issue in setting research agendas and making policy.

  – Evaluation provides a scientific basis for PS

  – It is very important to understand the links between evaluation and PS.

  – The role of (ex post) evaluation, especially in priority setting, could be the third theme of this project.

- The work could be done parallel to the case studies.

  – Method: Issues Paper & Small Workshop (Sep. 2008)

  – Deliverables: Best Practices in PS for research, including regarding the use of ex post evaluation

Results of the TIP/SFRI joint workshop on *PEER REVIEW* and *PRIORITY SETTING*

OECD 7 OCDE

Outline of the Workshop

- TIP & SFRI joint workshop on evaluation
  - 'Rethinking Evaluation in Science and Technology', 29-30 October, the French Ministry of Research in Paris.
  - 1st day: the role of peer review in the evaluation of research and innovation policies
  - 2nd day: the role of evaluation in priority setting/decision making for research and innovation policies

- This joint TIP-SFRI workshop aimed to:
  - Summarise the problems and issues;
  - Analyse approaches and solutions to these problems;
  - Identify good practices for peer review and priority setting

Issue Discussed on Peer Review 1

- ***Considering socioeconomic factors in evaluation:***
  - *How do we reflect socioeconomic and political priorities effectively and link these priorities to decision making in peer review processes?*
  - This is one of most important questions in policy evaluation

- How can we solve this problem?
  - To provide reviewers with pre-analysis of socioeconomic needs and priorities
  - To diversify the expertise of the reviewers
  - To have dual-level review committees:
    - eg NIH's 'Dual Review System' for grant applications
  - To use various evaluation methods with PR

- Do the basic assumptions of peer review hold in these circumstances?

Issue 2

- ***Interface of PR with other means of judgment:***
  - *How do we use indicators effectively in order to enhance the objectivity of evaluation result?*
  - *How can we combine peer review with quantitative and qualitative methods for evidence-based policy?*
- There are few cases where peer review is the *only* method used in policy and programme evaluation
  - Many indicators are based on past peer review judgements made for other purposes eg citations, grants awarded, prizes/esteem
  - Also various methods used in ATP evaluation
    - Expert judgment, survey, case study, sociometric/social network analysis, bibliometrics, historical tracing etc
  - US DOE uses various methods to obtain information on programme effectiveness and realized benefits that cannot be provided using the peer review method

Issue 3

- ***Cost efficiency of peer review:***
  - *How do we enhance the cost efficiency of the various parts of the peer review process?*

- Benefits of evaluation should outweigh the costs of it
  - Cost of PR is easily underestimated because usually incurred as an implicit opportunity cost not an explicit payment
  - NSF reduces the number of final proposals by comparing the results of "mail review" at the first stage evaluation with the ones of "panel review" at the second stage evaluation
  - SRP of NIH saves evaluation costs by concentrating on only the quality proposals that rank 50% and above
  - Recently, various types of alternative methods have been employed with the help of various internet supported tools
    - Eg NSF's Fast Lane System, widespread elctronic submission

Issue 4

- ***International frame of reference:***
  - *How can we establish an effective international frame of reference for peer review?*

- In a global research and innovation system standards or approaches should be defined internationally
  - An international frame of reference is increasingly used as the standard for peer review
  - Trade-off between independence and contextual knowledge
  - The internationalisation of science itself requires international reference points in measuring outcomes
  - But potentially reduces its own benefits as cooperation reduces probability of finding true independents

Issue 5

- ***Managing conflicts of interest:***
  - *How do we manage conflicts of interest in the PR process?*
  - *What is the best way to reach the final decision effectively?*
- Evaluator's decision is potentially affected by personal relationships which could prevent an impartial and objective evaluation
  - It is however nearly impossible to nominate expert review panels who have absolutely no interest
- Conflicts of interests also could occur between an evaluation manager and a reviewer
  - Who is responsible for the evaluation results?
  - Is it a manager or a reviewer who should make the final resource allocation decision?

Issue 6

- *"Open Evaluation" in the internet age:*
  - *What opportunities does the internet give us for improved and enhanced peer review?*
  - *Could an internet based "open evaluation" tool organized by the scientific community be an alternative to the classical approach?*
  - *Is evaluation possible without peer review panels?*

- Internet provides opportunities for advanced evaluation as well as new means and modes of communication
  - Internet conveys all kinds of useful information and data analysis tools in real time

- Internet makes possible a new style of PR
  - Internet-based "open evaluation" can secure additional evaluators around the world without a boundary and could be a very powerful tool to detect data fabrication
  - Interactive open access publishing of *JACP* gives a good example

Issue 7

- ***Crisis of confidence:***
  - *What are key factors causing the crisis of confidence?*
  - *How do we resolve crisis?*
- PR is affected by different factors, which have nothing to do with the evaluation object
  - "Matthew effect"
  - Cronyism, Informal cartel or personal connection
  - Favoritism
  - Discrimination against emerging or interdisciplinary fields
  - "Old-Boys-Network"
  - Conservatism: "Is this research successful?"
  - Ethical issues: Fraud, Plagiarism, Fabrication etc.
- But at political level the main threat is a perception that peer review creates perverse incentives away from desired goals such as working with business

Issue 8

- ***Peer review for policy, programme and/or Public Research Organisations:***
  - *What type of PR is fit for the evaluation of policy, programme, or PROs?*
  - *Is PR a relevant tool for evaluating research institutions?*
- What type of the peer review is appropriate for higher level decision making at programme, policy, or institute level
  - For example, following Bozeman, PR could be classified into a few categories based on the level of its impact on the final decision making:
    - pre-emptive peer review; traditional peer review; and ancillary peer review etc
- Some evidence that contractualisation of institute management has caused convergence with programme evaluation

Issue 9

- ***Good practices in peer review:***
  - *What constitutes good practice in peer review?*
  - *What policy recommendations could be made for better evaluation of policy, programme, or PROs?*

- A number of recommendations and alternatives have been suggested for the improvement of peer review
  - Ensure that experts declare their interest
  - Restrict the number of evaluations on which panel members serve
  - Broaden the panel as much as possible
  - Publicise the area of expertise of a particular panel member
  - Appoint the chair from among previous panel members
  - Seek experts from outside the geographical area where the programme is being carried out
  - Use various techniques such as "remote reviewer participation"

Bottom Line of the 1ˢᵗ Day (Peer Review)

- The PR process remains a fundamental mechanism for both ex ante and ex post evaluation.

- Solutions to improve PR:
  - more transparent process
  - clear objectives and guidelines
  - using different tools
  - using a variety of indicators.

- While indicators can strengthen and inform judgements, they are not a form of judgement in themselves.

- There is a need to improve the internationalisation of PR because of increased international collaboration.
  - There is a need for a taxonomy of the internationalisation of PR.

- One size does not fit all.
  - Better understanding of the design requirements for PR is needed

Bottom Line of the 2nd Day (Priority Setting)

- Although PS and evaluation interact in policy making, they remain two distinct dimensions of policy making.

- PS has become more complex and involves more actors using different approaches and methodologies.

- Expert opinion continues to predominate in the types of evaluation to make policy decisions and set priorities.

- Improving the process of PS requires:
    - political "buy-in" from the different stakeholders;
    - commitment to invest in resources and develop skills;
    - data to monitor policy or programme effectiveness.

- The process of PS itself could be the subject of evaluation to identify structural weaknesses as well as best practices.

- The interest of the international community is essential to develop the use of ex ante evaluation in PS.

FRAMEWORKS and *GUIDELINES*
for case studies

General Guidelines

- The case study should/could:
  - Include a short description of the reasons why a specific case study was selected and the methodology used.
  - Include a brief description of the development of the evaluation.
  - To the extent possible, follow the analytical framework proposed.
  - Highlight the important themes in the OECD evaluation activity.
  - Present lessons and suggestions for future development.
  - Consider the various opinions from experts and stakeholders.
- Countries may choose the methodology they wish.
- The indicators used in the case study should be clearly highlighted and listed in an annex.
- The case study could be limited to within 20 pages.

Analytical Framework

- To ensure some degree of *comparability,* the Secretariat proposes an over-arching meta-evaluation framework
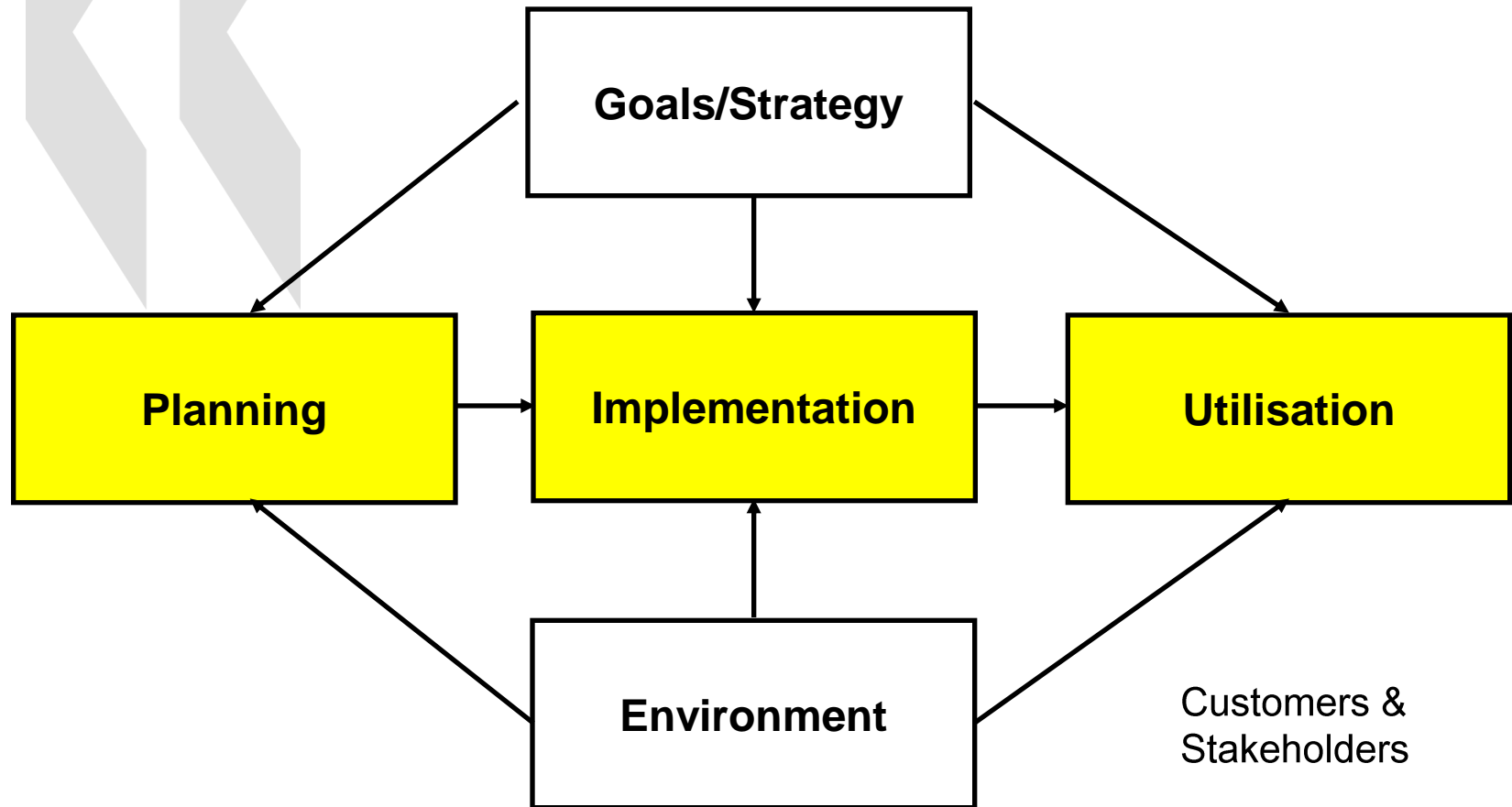
Elements & Related Key Questions

- ***Environment:*** the related rules and acts, the IT infrastructure, information systems, culture, education for stakeholders, openness to foreign evaluators etc.
  - What is the regulatory or legal basis for the evaluation of public R&D?
  - What are the institutional frameworks that support evaluation?
  - Is there an information system for collecting and analysing information on evaluations, and if so, how was it designed?
  - What education or training system is available for the evaluation of public R&D?

- ***Goals & Strategy:*** the philosophy, purpose, principle, and scope of the evaluation etc.
  - What is the main purpose of the evaluation?
  - What strategies are used to attain the goals of the evaluation?
  - What is the function and role of evaluation in public R&D and innovation policies?
  - How do the stakeholders interact in setting the goals and scope of the evaluation?

Elements & Related Key Questions

- *Planning:* the resources (the players, money, time, organisations) for evaluation, the design and planning mechanisms etc.

  – How large is the budget (and time) allocated for evaluation activities?

  – Are the evaluator's activities aligned with the mission of the evaluation system under study?

  – Are the detailed plans for evaluation set up systematically and strategically?

  – Who participates in the design of the evaluation?

  – What are the most important factors to be considered in planning the evaluation?

  – Is the evaluation plan known in detail to the stakeholders before it is implemented?

  – How are the indicators used for evaluation selected and agreed upon?

  – Is international benchmarking used in the planning of the evaluation, and if so, how is it used in practice?

Elements & Related Key Questions

- ***Implementation:*** the process and methodology of evaluation, the role and activities of decision makers, programme managers, external experts, and other stakeholders.

  – What are respective roles of the evaluation commissioners in managing the evaluation?

  – What are the roles and functions of external experts and other stakeholders?

  – What are the key processes in the evaluation?

  – What methodologies are used in the evaluation?

  – What indicators are used and how is the evaluation measured?

  – How is the evaluation committee (or panel) organised?

  – How is expert review or peer review used in the evaluation?

  – How are the materials and information analysed and provided to the evaluators?

  – Do evaluators and evaluatees interact during the evaluation process and if so, how?

  – How is the final decision on the evaluation made?

Elements & Related Key Questions

- ***Utilisation:*** the feedback mechanism (that is, the use of evaluation results in priority setting, budget allocation, and policy decision making), the system for monitoring the activities of stakeholders

  – Who are the primary and secondary users of the evaluation findings?

  – Do the evaluatees readily use the evaluation findings to improve policies and programmes?

  – How are the results of the evaluation used in priority setting?

  – How are the results of evaluation reflected in the budget co-ordination & allocation process?

  – How effective is the evaluation in influencing decision making and improving policies?

  – How does the evaluation contribute to a stronger planning ability of the manager?

  – Is a meta-evaluation practiced subsequent to an evaluation?

  – How effective is the meta-evaluation tool for improving evaluations?

NEXT STEPS of the project

Proposed Schedule

- **February – May 2008**
  - Carry out case studies in volunteer countries
  - Finalise the draft of Case study by 16 May 2008

- **June – August 2008**
  - Discuss drafts of the case studies at TIP & Revise the drafts
  - Synthesise the case studies and engage in further study as needed
  - Prepare draft synthesis report

- **September 2008**

  - Joint workshop on *Case Studies* and *Priority Setting*

- **October 2008**

  - Draft of the final synthesis report presented to CSTP

- **December 2008 – Early 2009**
  - Discussion of final report at TIP and declassification for publication
  - Publication of report

Next Steps of Evaluation Project

Issues	Next Steps	Deliverables
Case Studies	▪In progress ▪Possible TIP/SFRI Joint workshop (September 2008) ▪To be finished by the end of 2008	Best Practices and Recommendations
Expert Review (Peer Review)	▪Revision of summary report ▪Collection of additional material ▪Extended literature review	Best Practices and Recommendations
Priority Setting	▪Revision of summary report ▪Collection of additional material ▪Extended literature review ▪Possible TIP/SFRI Joint workshop (September 2008)	Best Practices and Recommendations

OECD ((29 OCDE

You are invited to discuss and comment on this presentation