

# Analysis of the Results of Intermediate and Ex-post Evaluations at NEDO

3 November 2006

AEA Conference, Hilton Hotel, Portland, OR

**Momoko OKADA**

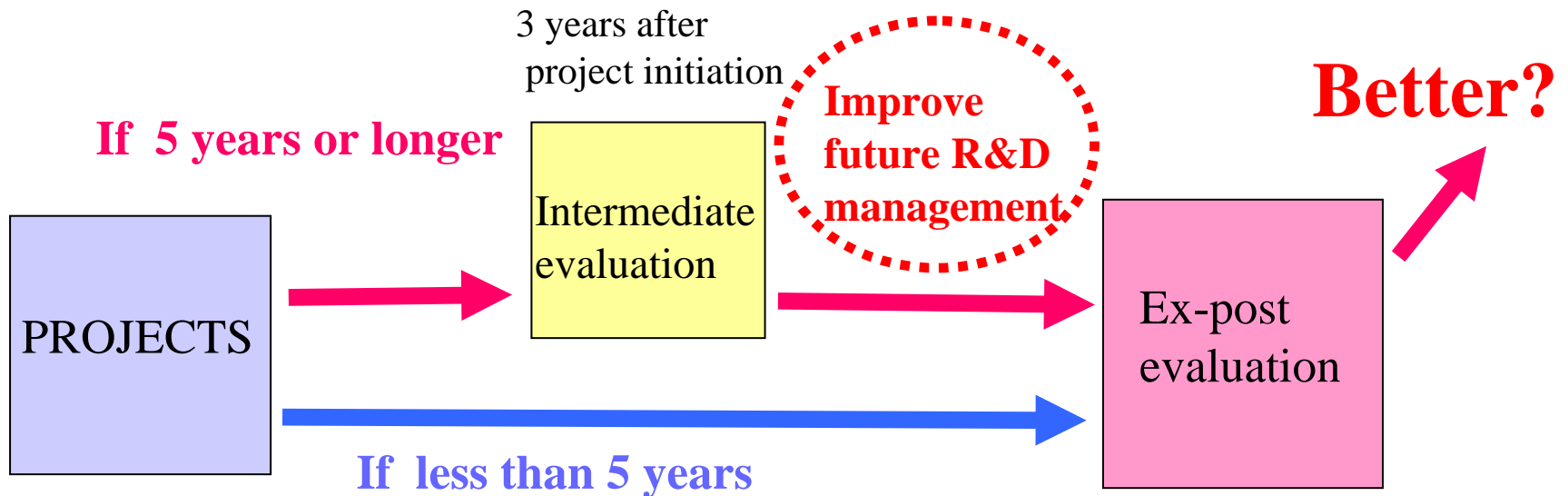
R&D Project Evaluation Department  
New Energy and Industrial Technology  
Development Organization(NEDO)

# Contents

- Objectives
- Target Projects
- Standard Evaluation Categories and Criteria
- Rating Method
- Results and Analysis
- Conclusion

# Objectives

- ◆ When conducting intermediate evaluation
  - Examine future implementation
  - Reflect results in the budget requests for subsequent fiscal years



# Objectives

- ◆ Our hypothesis
- Intermediate evaluation yields better results in the ex-post evaluation stage
  - ➔ study hypothesis

# Target Projects

- ◆ Projects for which ex-post evaluation was conducted in FY2003~2005

50 projects — Only ex-post evaluation conducted

24 projects — Both intermediate and ex-post evaluation conducted

Year	Target projects		
	Intermediate evaluation conducted in the past	No evaluation conducted in the past	Ex-post evaluation total
FY2003	6	20	29
FY2004	11	19	30
FY2005	7	8	15
FY2003~2005	24	50	74

# Standard Evaluation Categories and Criteria

► The same evaluation categories and criteria are used in both intermediate and ex-post evaluations.

<p><b>1. Purpose and Strategy</b></p> <ul style="list-style-type: none"><li>(1) Validity as a NEDO project</li><li>(2) Validity of objectives</li></ul>	<p><b>3. R&amp;D Achievements</b></p> <ul style="list-style-type: none"><li>(1) Success at achieving objectives</li><li>(2) Significance of research results</li><li>(3) Patent acquisition</li><li>(4) Publication and dissemination</li></ul>
<p><b>2. R&amp;D Management</b></p> <ul style="list-style-type: none"><li>(1) Validity of R&amp;D goals</li><li>(2) Validity of R&amp;D plan</li><li>(3) Validity of project team formation</li><li>(5) Response to change in situation</li></ul>	<p><b>4. Prospects For Practical Application and Other Impacts</b></p> <ul style="list-style-type: none"><li>(1) Feasibility of practical application</li><li>(2) Scenario for commercial application</li><li>(3) Ripple effect</li></ul>

# Standard Rating Method

## ◆ Intermediate & ex-post evaluations

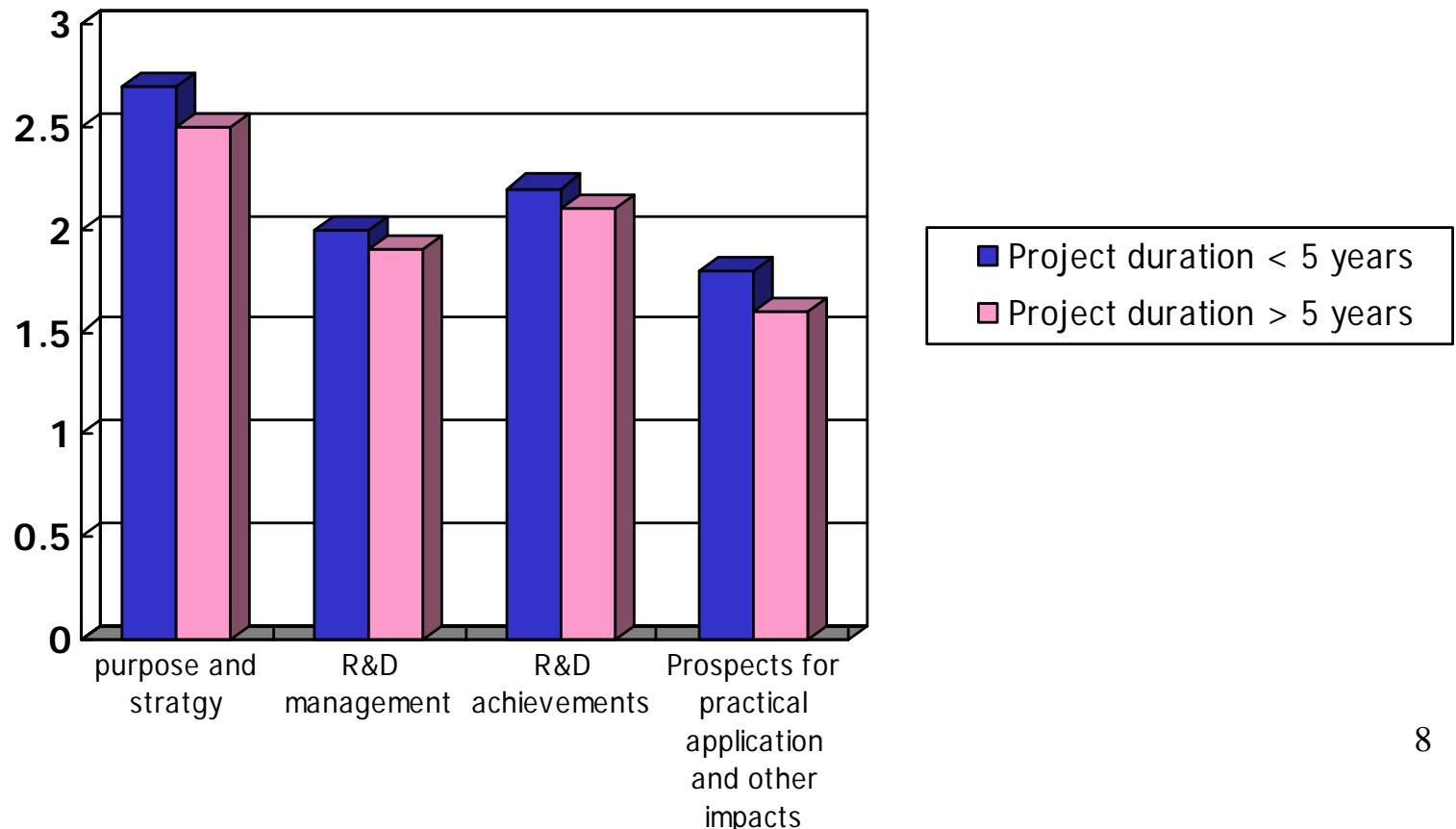
### - R&D Evaluation Methodology (Rating Method)

1. Each evaluator grades (A, B, C or D) for each evaluation item.  
A: Excellent, B: Good, C: Acceptable, D: Not acceptable
2. A numerical average of the rating is used as an indicator.  
A=3, B=2, C=1, D=0

Items for evaluation	Average	Rating					
		C	C	B	B	B	B
<b>1. Purpose and strategy</b>	<b>1.7</b>	C	C	B	B	B	B
<b>2. Project management</b>	<b>1.2</b>	C	C	C	C	C	B
<b>3. R&amp;D achievement</b>	<b>1.3</b>	C	C	C	C	B	B
<b>4. Prospects for practical application and other impacts</b>	<b>0.5</b>	D	D	D	C	C	C

# Results and Analysis

- ◆ Comparison of ratings of ex-post evaluations for projects that underwent intermediate evaluations and projects that did not undergo intermediate evaluations





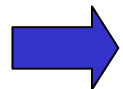
# Results and Analysis

- ◆ Findings
  - Contrary to our hypothesis, ex-post evaluation results for long-term project were slightly lower than those of short-term R&D projects.
  - “Prospects for practical application and other impacts” received the lowest score in both groups.
  - There was a discrepancy in results from the “Prospects for practical application and other impacts” category.

# Results and Analysis

## ◆ Possible explanations

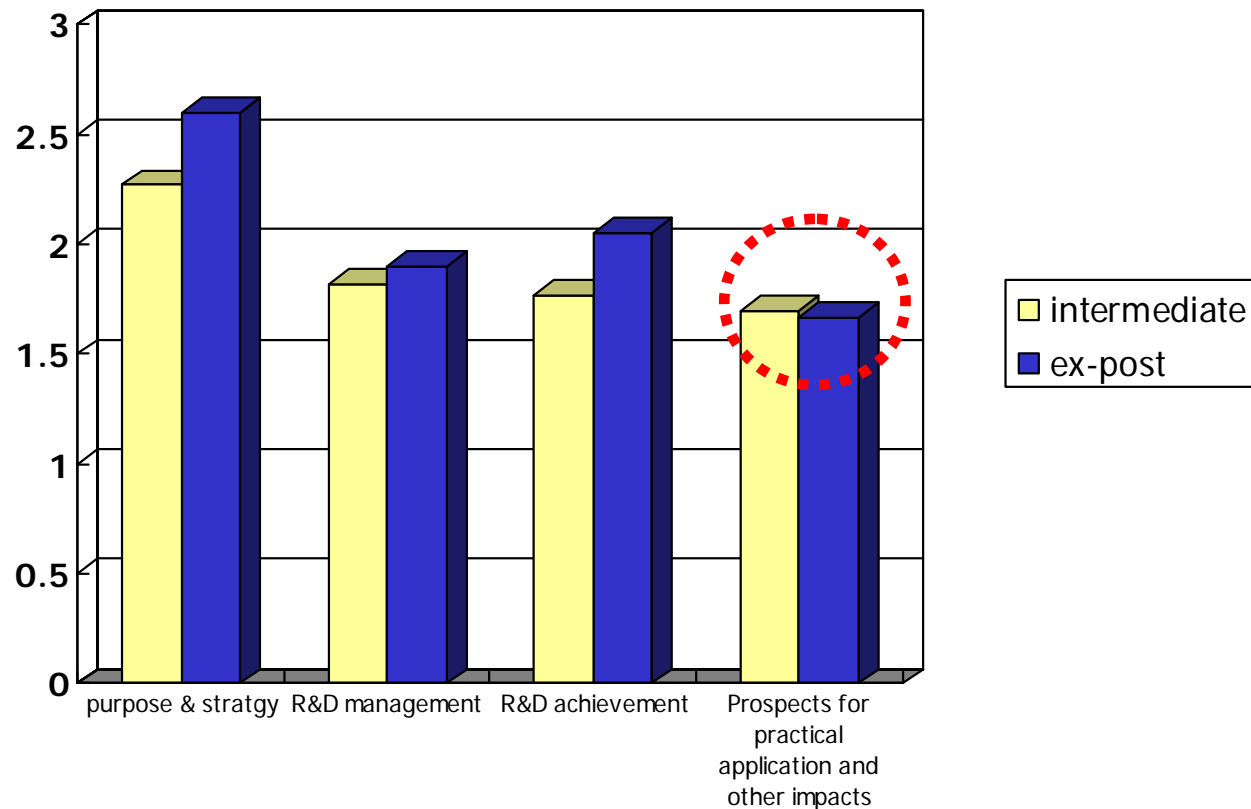
- Relative to long-term projects, short-term projects are more focused on the immediate application of their achievements, putting more emphasis on practicality and purpose of R&D.
- Long term projects tend to involve high-risk, and the rating results could be lower without intermediate evaluation



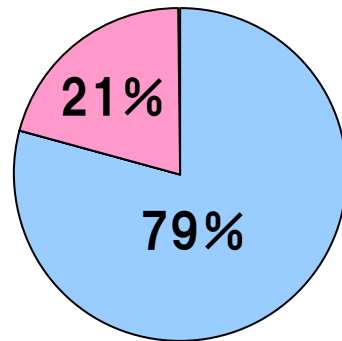
Study the 24 projects that received intermediate evaluation in greater detail

# Results and Analysis

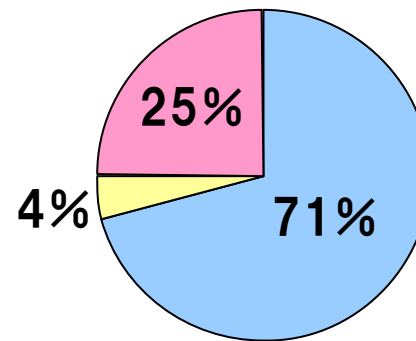
- ◆ Comparison of ex-post evaluation scores and intermediate evaluation scores



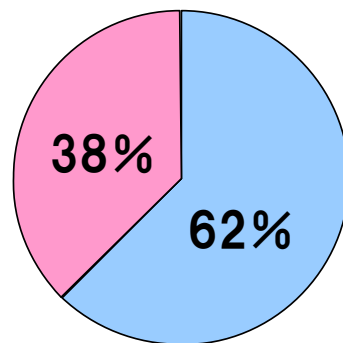
# Results and Analysis



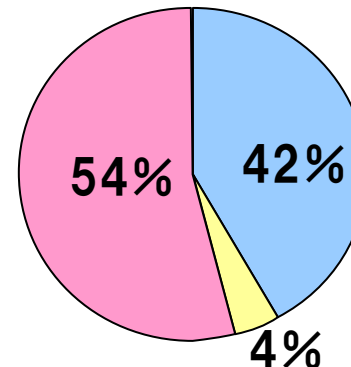
**Purpose and strategy**



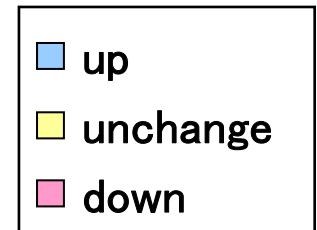
**R&D achievement**



**R&D management**



**Prospects for practical application and other impacts**



# Results and Analysis

## ◆ Findings

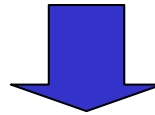
- Scores from ex-post evaluations are all higher than intermediate evaluations, except for in the category, “Prospects for practical applications and other impacts.”
- The score of “Prospects for practical application and other impacts” usually remains unchanged.

Scores are higher in general with the exception of the category, “Prospects for practical application and other impacts.”

# Results and Analysis

## ◆ Possible explanations

- The intermediate evaluations have contributed to the improvement of the performances of the projects at the ex-post evaluation, except for “Prospects for practical applications and other impacts.”
- Regarding “Prospects for practical applications and other impacts”, it is uncertain whether the results of intermediate evaluations reflected the real situation or not.

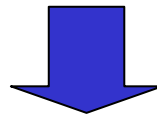


Evaluation questions on “Prospects for practical applications and other impacts” might be difficult to answer at the intermediate evaluation?

# Results and Analysis

## ◆ Possible explanations (cont.)

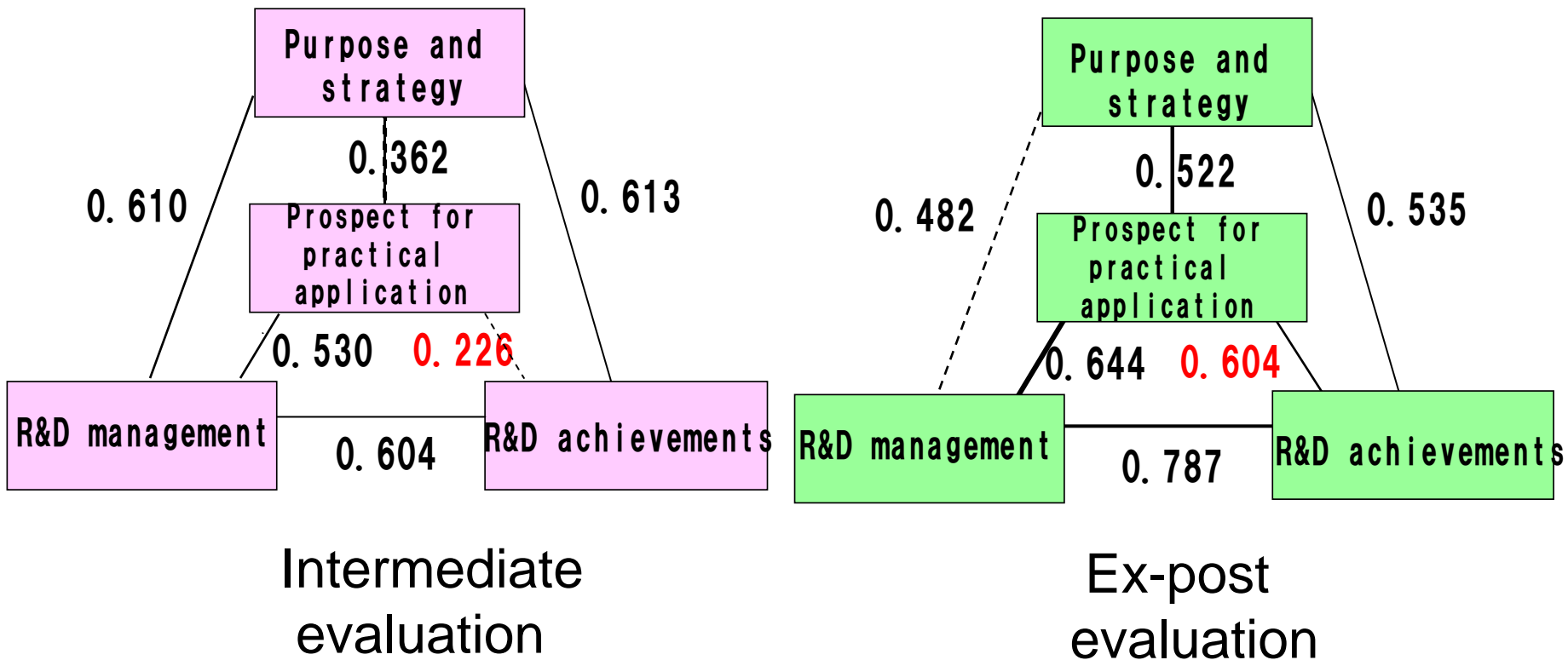
- Evaluations of “Purpose and strategy”, “R&D management” and “R&D achievements” reflect the current situation of the projects, while only the evaluation of “Prospects for practical application and other impacts” is based on future expectations
- It is likely that there are correlations among the three evaluation categories: “Purpose and strategy”, “R&D management” and “R&D achievements”, and evaluators might benefit from such correlations when they score these evaluation categories. However, evaluators might not be able to consider the scores of the other three evaluation categories when they score the evaluation category of “Prospects for practical application and other impacts” .



➤ Study the correlations among evaluation categories

# Results and Analysis

## ◆ Correlations among evaluation categories





# Results and Analysis

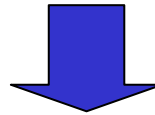
## ◆ Findings

- At intermediate evaluation, “Prospect for practical application and other impacts” is relatively uncorrelated to other evaluation categories.
- At ex-post evaluation, the correlations among “Prospect for practical application and other impacts” and other evaluation items are stronger than at intermediate evaluation.
- In particular, correlation between “R&D achievements” and “Prospect for practical application and other impacts” is very weak at intermediate evaluation but strengthened at ex-post evaluation.

# Results and Analysis

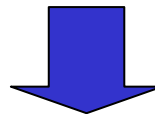
## ◆ Possible explanations

- At ex-post evaluation, results of “R&D achievements” become clearer and “Prospects for practical application and other impacts” are recognized more precisely



**Appropriate evaluation at ex-post evaluation**

- As NEDO promotes “medium and long term, high-risk” projects, it is difficult to evaluate “Prospects for practical application and other impacts” precisely at the stage of intermediate evaluation because evaluators can only see half of the “R&D achievement” and evaluation questions might be difficult at that stage.



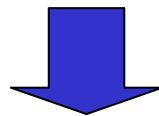
**Less accurate evaluation at intermediate evaluation**

# Conclusion

- Intermediate evaluations yield better results at the ex-post evaluation stage. This is true for “Purpose and strategy”, “R&D management” and “R&D achievements”.
- However, it is difficult to evaluate “Prospects for practical application and other impacts” at the intermediate evaluation. As a result, intermediate evaluations do not always foster improvement in this category by the time of ex-post evaluations.

# Conclusion (cont.)

- Evaluations of “R&D achievements” involve more uncertain factors at the intermediate stage than at the time of ex-post evaluation.
- If evaluation of “R&D achievements” becomes clearer at the time of intermediate evaluation, and if the evaluation questions for “Prospects for practical application and other impacts” become clearer, correlations between “R&D achievements” and “Prospects for practical application and other impacts” might be improved such as in the case of ex-post evaluation, and thus evaluators might be able to estimate “Prospects for practical application and other impacts” better.



We should introduce new evaluation questions in which “R&D results” and “Prospects for practical applications and other impacts” can be evaluated more precisely at the intermediate evaluation.

**Thank you!**