EFFORTS TO IMPROVE THE RATING SYSTEM

JBIC analyzes rating results based on the past ex-post evaluations. By utilizing and improving the rating system, JBIC is striving to improve aid effectiveness of development projects. Below appears an example of results in our analysis.

Outline and Objectives

Based on a recommendation made in the current and last fiscal year by the Yen Loan Evaluation Expert Committee*, we analyzed trends in ratings according to the results of ex-post evaluations conducted between FY2001 and FY2006. (The total number of subject for this analysis was 324 projects, the said number for part of the analysis was 217.) Specifically, a study and numerical analysis was tentatively applied to the distribution of the results of rating, the relationship between ratings and project scale, trends in efficiency (duration) and sustainability for the results of rating, and make comparison of the results of rating with utilizing scoring methods, which is different from JBIC's current rating system. As the analysis results shown below are tentative, we will continue to strive to analyze such matters in an effort to further improve the rating system.

* In FY2006 the name was changed from the "Feedback Committee on Ex-Post Evaluation of ODA Loan Project "

1. What Is the Distribution of the Results of Rating in the Current Rating System?

Rating Score Distributions (217 projects in total)

	Relevance		Effectiveness		Efficiency		Sustainability	
a	210	96.8%	155	71.4%	29	13.4%	92	42.4%
b	7	3.2%	52	24.0%	156	71.9%	109	50.2%
c	0	0.0%	10	4.6%	31	14.3%	16	7.4%
NA					1	0.4%		

	Overall Rating	
A	74	34.1%
B	84	38.7%
C	39	18.0%
D	19	8.8%
NA	1	0.5%

* Out of 217 projects, none had a relevance rating of "c," and only very few (7 cases) had a rating of "b."
* As many as 71.9% of projects (156 cases out of 217) had an efficiency rating of "b." At the same time, 155 projects had an effectiveness rating of "a" (71.4%). While there is still room to improve project implementation from the standpoint of efficiency, overall, project effectiveness was judged to be high.
* Out of 217 projects, 109 (50.2%) had a sustainability of "b." If combined with the 92 projects which were provided "a" ratings, "a" and "b" combined accounted for 92.6% of all projects.

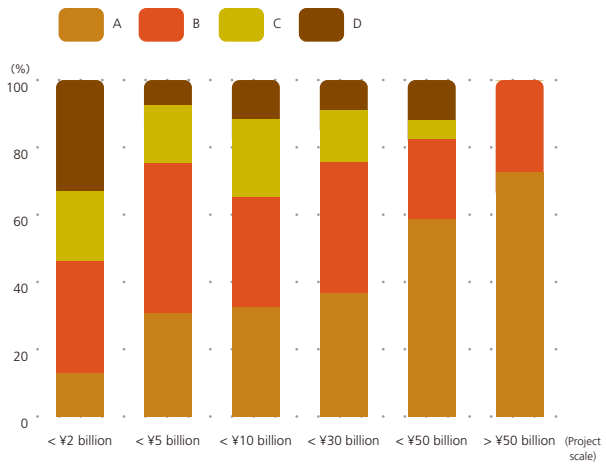
Five countries—namely, India, Indonesia, the Philippines, Thailand, and China—accounted for a vast number of the projects. Based on the analysis of the 139 projects conducted in these countries, the following facts were revealed:

- In India, Thailand, and the Philippines, efficiency (duration) ratings of many projects were provided as "c," but project costs were held down to within planned figures.
- In Thailand, effectiveness rating regarding 13 out of 14 projects (92.9%) was provided as "a."
- Of the projects undertaken in the five main countries, only 6 projects (4.3% of cases) of sustainability rating were provided as "c," and in Thailand, none had such a low rating.

2. What Is the Relationship Between Project Scale and the Results of Rating?

All 324 projects as the subject of this analysis were grouped according to project scale. By comparing the groups, it became apparent that the larger the project, the more likely it was to have an overall "A" rating. Conversely, looking at the composition of projects with "D" ratings, projects that were in the smallest scale group had the highest proportion of "D" ratings (33.3%), whereas none of the largest-scale group had a "D" rating. As a result, it could be possible to assume there is the tendency that larger projects would attain positive evaluations because of some sort of contributing factors.

Project scale and the results of rating



3. What Was the Contributing Factor for Extended Project Durations?

Looking at overall project duration, out of 320 projects (except for four projects whose data was not available), the average duration was 91.8 months, which amounts to 175% of the initially planned duration. Based on that result, when we looked at the reasons for the delays described in the ex-post evaluation reports, "delays in procurement" (21.8%) were the highest proportion followed, in descending order, by "revisions and changes to plans" (16.8%), and "delayed procedures, negotiations, and adjustments" (11.9%).

4. Are Sustainability Ratings Being Provided Arbitrarily?

Evaluating sustainability of projects in the ex-post evaluation process, it should be analyzed regarding executing agencies with respect to four criteria, namely 1) technical capability, 2) system and organization, 3) financial status, and 4) operation and maintenance. In the evaluation report, these criteria are assessed as "no problem, "there are minor problems," or "there are problems." Based on the analysis of the report, it was found out that there are several projects where the sustainability was judged to have problems in only one of the four criteria, however, it was assigned a rating of "C." On the other hand, although there were projects deemed to have problems in all four areas, it was judged sustainability rating as "B".

From these findings, it appears that evaluators may be applying arbitrary standards to the sustainability rating. In other words, because the standards for assigning this rating are rather vague, there is room for arbitrariness on the part of the evaluator. Therefore, it should be necessary to establish a set of clear standards by which to assess sustainability.

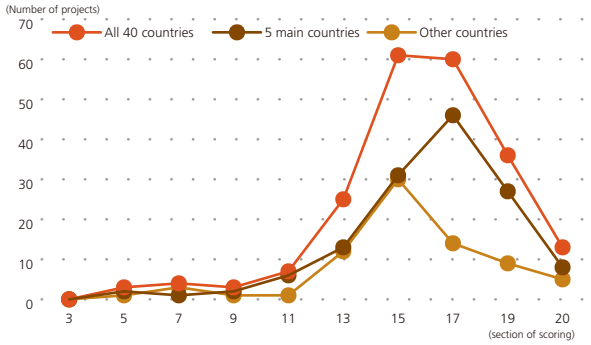
Patterns of Problems Arising with Respect to Sustainability
(*based on 319 cases for which sustainability evaluation was implemented based on the four criteria)

Classification	Number of cases	Rating		
		a	b	c
No problems in any criteria	163	116 71.2%	47 28.8%	0 0.0%
A problem exists in one of the four criteria	80	26 32.5%	47 58.8%	7 8.7%
A problem exists in two of the four criteria	44	6 13.6%	29 65.9%	9 20.5%
A problem exists in three of the four criteria	22	3 13.6%	12 54.6%	7 31.8%
A problem exists in all four criteria	10	0 0.0%	7 70.0%	3 30.0%
Total	319*	151	142	26

5. What Is the Distribution According to Scoring Method?

A scoring method was conducted by weighting the current rating criteria. It was set an assumption that there was a total score of 20 points for each project and the allocation of scores was decided as follows: 4 for relevance, 6 for effectiveness, 6 for efficiency (3 for duration and 3 for project cost), and 4 for sustainability. Then, the distribution of total score for all 40 countries, and the five major countries was determined as shown below.

Distribution of Total Score for 5 Major Countries and All 40 Countries



The analysis indicates that the average score for the five major countries was higher than that for all 40 countries, and the standard deviation was slightly lower. From the results it could assume that overall the relative evaluations for the five major countries are higher than those for the other countries. The data also shows that projects in the five major countries were implemented more efficiently than projects in the other countries, as duration score averaged 1.36 out of a total possible 3 points for all 40 countries, whereas the five main countries had an average duration score of 1.5. This seems to indicate there might be possibility that executing agencies and their governments in five major countries were more capable of implementing projects efficiently comparing with other countries.

Further, the distributions were divided into three segments—high, medium, and low—and the typical trends within those segments were then analyzed.

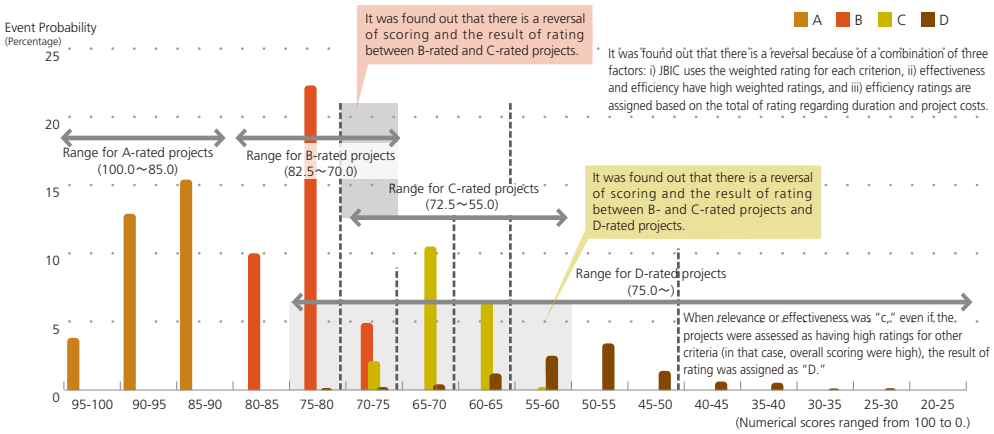
Projects Assigned Average Rating (Segment in Medium)
It was found out that one of the most common patterns were as follows for both all 40 countries and the five major countries: "a" for relevance, "a" for effectiveness, "b" for efficiency, and "b" for sustainability. However, among projects with overall ratings of "A" or "C," we find projects that had average scores, thus indicating there is a reversal between the results of ratings and scoring. In addition, among projects with average ratings, most of the projects assigned a "b" rating for efficiency were rated "c" for duration.

Projects Assigned Lower Ratings (Segment in Low)
More than half of the projects that had been assigned lower rating were assessed poorly because they had an effectiveness of "c" (10 out of 18 cases, or 56%). In addition, looking at projects with "D" ratings by sector, most were for agriculture (5 projects), and three each were in irrigation, flood control, and power plants.

Projects Assigned Higher Ratings (High Grade)
Projects assigned higher ratings were not confirmed any peculiar trends based on sector or country. However, among these projects, two constituted two-step-loan (TSL) projects, and all TSL projects had scores of 18 or higher and a rating of "A". Based on the result of this analysis it could be assumed that TSL projects might be evaluated too leniently rather than these are relatively well implemented.

6. Why Do the Result of Rating and Scoring Sometimes Not Correspond?

Under the current rating system, if relevance or effectiveness was "c", then the overall rating was automatically "D," which means that the other rating criteria were considered meaningless. For that reason, projects with a "D" rating have a wide range of scores, and it was found out that there are projects with "D" rating obtained scoring of 70-80. Moreover, because effectiveness and efficiency have high weighted scores, the results of rating and scoring could not correspond.



Conclusion

The above analysis was a tentative attempt to improve the rating system, and it is still in the process of being analyzed. Based on the results so far it was found out that there are problems in respect of the current rating system as indicated in the table to the right. To establish a system whereby well performed projects would receive good ratings, JBIC will continue to strive to analyze and study the matter so as to continuously improve the rating system.

Problem Category	Description
A) Problems of the rating system itself	•The majority of efficiency (duration) ratings were b or c. •The trend is for larger projects to have higher ratings.
B) Problems regarding unclearness of various criteria for evaluation standards and a degree of arbitrariness in the evaluation process	•Most relevance ratings were "a" •Most sustainability ratings were "a" or "b" (standards for the sustainability criterion were inconsistent or not yet established.)
C) Problems regarding the way of project implementation that was confirmed through ratings	•Based on the result of past implemented evaluation and rating, establishing appropriate weights and cutoff rates to minimize the inconsistencies between the results of scoring and rating.