

Third-party Q/C in 2018

Analysis report

The third-party quality check for
JICA ex-post evaluation (internal)

KPMG AZSA LLC

Overall of the targeted projects of third-party Q/C in 2018

- Number of targeted projects : 59 projects (Projects targeted for internal ex-post evaluation in FY2016 whose results were fed back to the implementing agency/ executing agency after October 2017)
- Scheme: 44 projects of Technical Cooperation (TC) (including 8 TC for Development Planning (DP)) and 15 projects of Grant Aid (GA)
- Overall evaluation results: 4 projects of highly satisfactory, 31 projects of satisfactory, 11 projects of partially satisfactory and 3 projects of unsatisfactory

List of the targeted projects of third-party Q/C in 2018

No	Country	Internal Evaluator	Project title	Scheme
1	Indonesia	Indonesia	Standardization and Quality Control for Horticulture Products of Indonesia (Improvement of Thermal Treatment Technique against Fruit Flies on Fresh Mango)	TC
2	Indonesia	Indonesia	Capacity Building for Enhancement of the Geothermal Development	TC
3	Indonesia	Indonesia	Project for Capacity Development of Jakarta Comprehensive Flood Management	TC
4	Indonesia	Indonesia	The Project for PPP Network Enhancement	TC
5	Cambodia	Cambodia	The Legal and Judicial Development Project (Phase2)	TC
6	Cambodia	Cambodia	The Legal and Judicial Development Project (Phase3)	TC
7	Cambodia	Cambodia	The Project for Introduction of Clean Energy by Solar Electricity Generation System	GA
8	Cambodia	Cambodia	The Project for Improvement of Medical Equipment in National, Municipal and Provincial Referral Hospitals	GA
9	Viet Nam	Viet Nam	Project on Human Resource Development for Urban Water Supply Utilities in Central Region	TC
10	Thailand	Thailand	Land Readjustment Promotion Project	TC
11	Myanmar	Myanmar	Strengthening of Rehabilitation	TC
12	Myanmar	Myanmar	Small-scale Aquaculture Extension for Promotion of Livelihood of Rural Communities in Myanmar	TC
13	Philippines	Philippines	The project for study on improvement of the bridges through large scale earthquakes disaster mitigating measures	TC(DP)
14	Solomon Islands	Solomon Islands/ Papua New Guinea	Project for Strengthening of Malaria Control System Phase II	TC
15	Mongolia	Mongolia	Capacity Development Project for Air Pollution Control in Ulaanbaatar City	TC
16	Mongolia	Mongolia	Capacity Development for Promoting Foreign Direct Investment	TC

List of the targeted projects of third-party Q/C in 2018

No.	Country	Internal Evaluator	Project title	Scheme
17	Mongolia	Mongolia	Study on the Strategic Planning for water supply and sewerage sector in Ulaanbaatar City	TC(DP)
18	Mongolia	Mongolia	The Project for Improvement of Capacity of Fire Fighting Techniques and Equipment in Ulaanbaatar	GA
19	China	China	Sustainable Agricultural Technology Research and Development Phase 2	TC
20	China	China	Capacity Development Project for management plan of dam in China	TC
21	China	China	Human Resource Development Project for Seismic Engineering and Construction of Buildings, P.R.C	TC
22	China	China	Project on Forestry Human Resource Development in Western Region of China	TC
23	Bhutan	Bhutan	Strengthening of Quality of Vocational Education and Training Delivery	TC
24	Sri Lanka	Sri Lanka	Small Scale Dairy Farming Improvement through Genetic and Feeding Management Improvement	TC
25	Nepal	Nepal	Gender Mainstreaming and Social Inclusion Project	TC
26	Nepal	Nepal	The Project for Introduction of Clean Energy by Solar Electricity Generation System	GA
27	El Salvador	El Salvador	The Project for the Strengthening of Capacities for Rural Tourism Development in the Eastern Region of El Salvador	TC
28	Brazil	Brazil	The Jalapao Region Ecological Corridor Project	TC
29	Kosovo	Balkan	The Project for Improvement of Solid Waste Management	GA
30	Moldova	Middle East and Europe Department	The Project for Introduction of Clean Energy by Solar Electricity Generation System	GA
31	Iran	Iran	Project for Strengthening Environmental Management in Petroleum Industry in Persian Gulf and its Coastal Area	TC(DP)
32	Palestine	Palestine	The Project for Establishment of New Schools in the West Bank	GA
33	Kenya	Kenya	The Project for Strengthening of People Empowerment against HIV/AIDS in Kenya	TC
34	Kenya	Kenya	Project for Strengthening of People Empowerment against HIV/AIDS in Kenya (SPEAK) Phase 2	TC
35	Kenya	Kenya	Strengthening Management for Health in Nyanza Province	TC
35	Kenya	Kenya	The Project for the Reinforcement of Vaccine Storage in Kenya	GA
36	Nigeria	Nigeria	The Project for Rural Water Supply in Kano State	GA

List of the targeted projects of third-party Q/C in 2018

No.	Country	Internal Evaluator	Project title	Scheme
37	Nigeria	Nigeria	The Project for water Supply in Bauchi and Katsina States	GA
38	Rwanda	Rwanda	The Project for Capacity Building for Efficient Power System Development in Rwanda	TC
39	Ethiopia	Ethiopia	Participatory Forest Management Project in Belete-Gera Regional Forest Priority Area Phase 2	TC
40	Ethiopia	Ethiopia	The Ethiopian Water Technology Center Project Phase-3	TC
41	Ghana	Ghana	Project for Institutional Capacity Development of the Civil Service Training Centre	TC
42	Ghana	Ghana	The Project for the Study on Comprehensive Urban Development Plan for Greater Kumasi	TC(DP)
43	Djibouti	Djibouti/Ethiopia	The Project for Rural Water Supply in Southern Djibouti	GA
44	Djibouti	Djibouti/Ethiopia	The Project for Managing Digital Topographic Data in Djibouti City	TC(DP)
45	Tanzania	Tanzania	The Project for the Comprehensive Transport and Trade System Development Master Plan	TC(DP)
46	Zambia	Zambia	The Project for the Capacity Development for Rural Electrification	TC
47	Malawi	Malawi	The Project for Groundwater Development in Mwanza and Neno	GA
48	Senegal	Senegal	Project for Treatment of Sewage, Rainwater and Wastes in Kaolack City	TC(DP)
49	Benin	Benin/Cote d'Ivoire	Le Projet de Renforcement des capacités de l'École Normale d'Instituteurs à Djougou	GA
50	Indonesia	Indonesia	Project for Capacity Development of Wastewater Sector through reviewing the Wastewater Management Master Plan in DKI Jakarta	TC
51	Laos	Laos	The Project for Improvement of Transportation Capacity of Public Bus in Vientiane Capital	GA
52	Laos	Laos	Project for Improvement of Power Sector Management	TC
53	Mongolia	Mongolia	The Project for Freshwater Resources and Nature Conservation	GA
54	Fiji	Fiji	ICT for Human Development and Human Security Project	TC
55	Cuba	Cuba	Capacity Development on Groundwater Development and Management for Climate Change Adaptation	TC
56	Niger	Niger/Burkina Faso	Malaria Control Project	TC
57	Burkina Faso	Burkina Faso	Project of Support for Seeding Production Sector	TC
58	Burkina Faso	Burkina Faso	Project for Rural Development through Aquaculture	TC
59	Gambia	Senegal	The Project for Rural Water Supply (Phase III)	GA

Objectives and Background

JICA evaluates projects for mainly fulfilling two purposes: 1) improve the projects further through the PDCA cycle and 2) ensure accountability to stakeholders including Japanese nationals and people of the partner countries. Internal ex-post evaluation is conducted by JICA overseas office staff and regional departments in the headquarter. The targets of internal ex-post evaluation are the projects which cost 200 million yen or more and under one billion yen.

In order to achieve the objectives of evaluation more effectively and efficiently, the internal evaluation process is being improved by introducing self-assessment by internal evaluators and quality check by external third parties.

The objectives of this study is to examine internal evaluations by external third-party and to ensure the quality, objectivity, and impartiality of future internal evaluations as well as to disclose its outline to the public to strengthen the accountability.

Tools and methodology of meta-evaluation

1. External meta-evaluation is conducted through the desktop research. External third parties examine internal evaluation report and verify them based on a check sheet similar to the one developed for self-assessment.
2. External meta-evaluator ticks Yes/ Partly Yes/ No/ Not Applicable for each checklist item and comment the reasons on the check sheet.
3. Overall trends (e.g. average and characteristic results etc.) are analyzed based on the meta-evaluation of individual projects.
4. For analysis of the overall trends, the results of the meta-evaluation are calculated as score.
5. Information source for meta-evaluation is limited to the internal evaluation report in principle.
6. Meta-evaluation check sheet is aligned with the JICA Guidelines for Project Evaluation and Internal Evaluation Implementation Manual.

Methodology of scoring

① **Raw score :**

Score 2 points for “Yes”, 1 point for “Partly Yes”, 0 point for “No”, and no score for “Not Applicable”.

② **Standardized score :**

Standardized score = total of raw score ÷ (total number of checklist items – number of tick on “Not Applicable”)×2

Standardized score is between 0.0 to 1.0. As the score is closer 1, it means that there are more ticks on “Yes”, in other words, the internal evaluation’s quality is higher.

【Note : Calculation of dispersion and standard deviation】

$$\text{Dispersion : } \frac{1}{n} \sum_{n=1}^n (x_i - \bar{x})^2 \quad \text{Standard deviation : } \sigma = \sqrt{\frac{1}{n} \sum_{n=1}^n (x_i - \bar{x})^2}$$

n = total number of targeted projects、 x_i = average of standardized score、 \bar{x} = standardized score

The dispersion and standard deviation were calculated for the whole and for each evaluation criteria, and the variation in the meta-evaluation results of the target project was confirmed. The larger the dispersion and standard deviation, the larger the variation from the average score.

Items to be analyzed

1. Overall trends
2. Checklist items (①Relevance, ②Effectiveness/ Impact, ③Efficiency, ④Sustainability, ⑤Conclusion, Recommendations and Lessons Learned, ⑥Overall description)



Third-party Q/C of individual projects

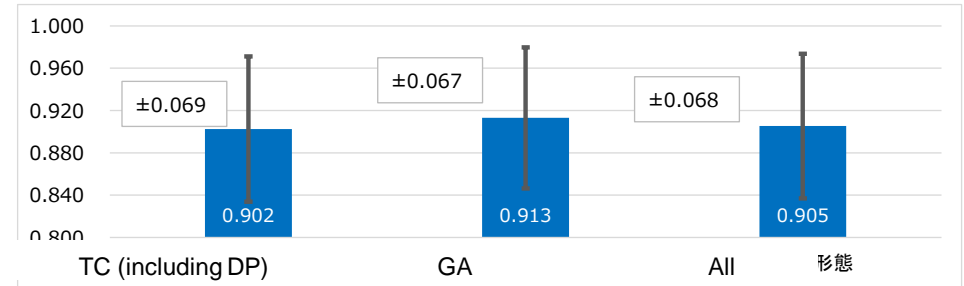


1. Analysis of overall trend

1. Analysis of overall trends

Average, dispersion and standard deviation for 59 targeted projects (standardized score)

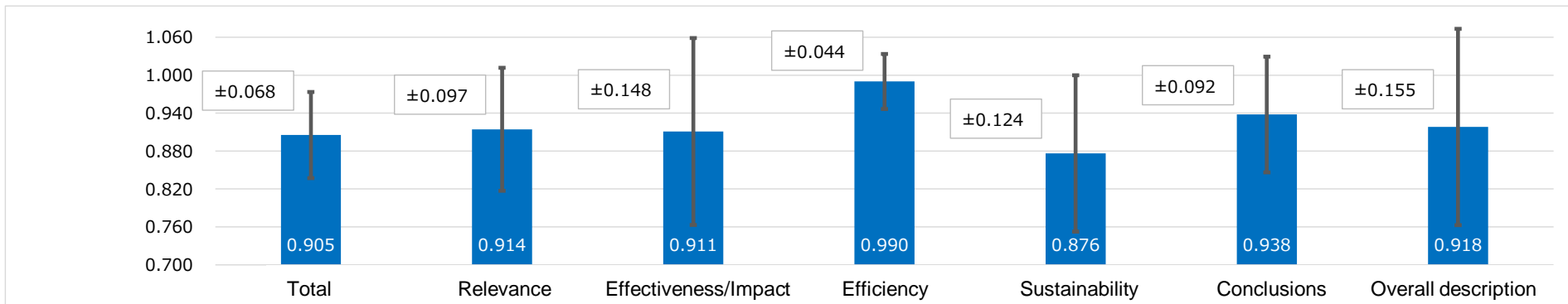
	TC	GA	All
Average	0.902	0.913	0.905
Dispersion	0.005	0.004	0.005
Standard deviation	0.069	0.067	0.068



Note : Bar graph shows average of standardized score and error bar (error range) shows standard deviation.

Average, dispersion and standard deviation for each criteria of 59 targeted projects (standardized score)

	Total	Relevance	Effectiveness/ Impact	Efficiency	Sustainability	Conclusion/ Recommendation and Lessons Learned	Overall description
Average	0.905	0.914	0.911	0.990	0.876	0.938	0.918
Dispersion	0.005	0.009	0.022	0.002	0.015	0.008	0.024
Standard deviation	0.068	0.097	0.148	0.044	0.124	0.092	0.155



Note : Bar graph shows average of standardized score and error bar (error range) shows standard deviation.

1. Analysis of overall trends

- As a result of the third-party Q/C, the average standardized score of all targeted 59 projects is 0.905. That means nearly 90% of checklist items are ticked “Yes” by meta-evaluation. Thus, it is possible to say that the quality of JICA’s internal ex-post evaluation and self-assessment is high.
- The average score of TC is higher than that of GA while there are not so much difference in terms of variation of them.
- Efficiency shows high average score, on the other hand those of Effectiveness/Impact and Sustainability are low.
- Effectiveness/Impact, Sustainability and Overall description show larger variation. As for Overall description, the influence in case that the score is deducted by third-party Q/C becomes relatively larger than the other criteria, and the variation tends to be large. It is because number of checklist items of Overall description is relatively small and many of them are ticked on “Not Applicable”.
- Efficiency shows high average score as well as small variation, therefore it can be said that the evaluation quality of this criteria is high.
- Effectiveness/Impact and Sustainability show low average score and large variation, means that there are relatively many difference between internal ex-post evaluators in terms of quality and accuracy.

1. Analysis of overall trends

This section shows the result of gap analysis between self-assessment and meta-evaluation. In more detail, the gap of raw score of 57 projects (targets of self-assessment), and the difference between criteria.

Raw score Analysis of targeted 59 projects

	(a)Self-assessment	(b)Meta-evaluation	Gap (b)-(a)
Total of raw score	2,547	2,362	-185
Average raw score of each project	44.7	41.4	-3.3

Note : Self-assessment were done for 57 projects of 59 third-party Q/C's targets. Therefore, the gap analysis between self-assessment and third-party Q/C are conducted for these 57 projects.

Gap analysis of checklist items between self-assessment and meta-evaluation

Type of gap	Number of ticks
Meta-evaluation = Self-assessment	1,140 (71.4%)
Meta-evaluation > Self-assessment	146 (9.1%)
Meta-evaluation < Self-assessment	163 (10.2%)
Number of items those are scored by meta-evaluation but ticked on "Not Applicable" by self-assessment.	147 (9.2%)

Note : The population of this analysis is calculated as 57 projects (targets of self-assessment) × 28 checklist items = 1,596 items

The 5 items are recognized to be ticked as same as third-party Q/C, because they are not ticked by self-evaluator.

Items those are ticked on "NA" by self-evaluator while are ticked on "Yes/Partly Yes/No" by meta-evaluator are included "Meta-evaluation > Self-assessment" as a type of gap.

1. Analysis of overall trends

Evaluation criteria etc.	Difference between self-assessment and meta-evaluation	Number of ticks (percentage to population)	Meta-evaluator's comments
Relevance (57projects × 5items = 285)	Meta-evaluation = Self-assessment	210 (73.7%)	<ul style="list-style-type: none"> ● The result of meta-evaluation is almost consistent with that of self-assessment. ● In many cases, self-assessors count score for checklist item No.4, however meta-evaluator tick on "NA".
	Meta-evaluation > Self-assessment	25 (8.8%)	
	Meta-evaluation < Self-assessment	20(7.0)	
	Ticked on "NA" by meta-evaluation	30(10.5%)	
Effectiveness/ Impact (57projects × 6items = 342)	Meta-evaluation = Self-assessment	266 (77.8%)	<ul style="list-style-type: none"> ● Among all checklist items, there are the largest number of items those scores of meta-evaluations are lower than self-assessments. ● This trend is same among checklist items No.6 to No.11.
	Meta-evaluation > Self-assessment	22(6.4%)	
	Meta-evaluation < Self-assessment	54(15.8%)	
	Ticked on "NA" by meta-evaluation	0	
Efficiency (TC41projects × 3items + GA16projects × 4items = 187)	Meta-evaluation = Self-assessment	134 (71.7%)	<ul style="list-style-type: none"> ● In many cases, self assessors count score for checklist items of TC's No.13, GA's No.12 and No.14, however meta-evaluator tick on "NA".
	Meta-evaluation > Self-assessment	9 (4.8%)	
	Meta-evaluation < Self-assessment	1 (0.5%)	
	Ticked on "NA" by meta-evaluation	43(23.0%)	
Sustainability (TC41projects × 7items + GA16projects × 6items = 383)	Meta-evaluation = Self-assessment	277(72.3%)	<ul style="list-style-type: none"> ● Following Effectiveness/Impact, Sustainability shows second largest number of items those scores of meta-evaluations are lower than self-assessments. ● This trend significantly occurs in TC's checklist No.19 and No.20.
	Meta-evaluation > Self-assessment	45 (11.7%)	
	Meta-evaluation < Self-assessment	58(15.2%)	
	Ticked on "NA" by meta-evaluation	3(0.8%)	

1. Analysis of overall trends

Evaluation criteria etc.	Difference between self-assessment and meta-evaluation	Number of ticks (percentage to population)	Meta-evaluator's comments
Conclusion/ Recommendations/ Lessons Learned (57projects × 4items = 228)	Meta-evaluation = Self-assessment	137 (60.0%)	● There are no checklist item of No.25 (appropriateness of no description of recommendation and lessons learned) in self-check sheet. The left results shows much percentage of "NA", because meta-evaluators tick on "NA" in many cases.
	Meta-evaluation > Self-assessment	19 (8.4%)	
	Meta-evaluation < Self-assessment	19 (8.4%)	
	Ticked on "NA" by meta-evaluation	53 (23.2%)	
Overall description (57projects × 3items = 171)	Meta-evaluation = Self-assessment	116 (67.8)	● In many cases, self-assessors tick on "NA" for checklist No.26, however meta-evaluators count scores for this item.
	Meta-evaluation > Self-assessment	26 (15.2%)	
	Meta-evaluation < Self-assessment	11 (6.4%)	
	Ticked on "NA" by meta-evaluation	18 (10.6%)	

1. Analysis of overall trends

- The difference of average raw score between self-assessment and meta-evaluation is about 3.3points each project, and self-assessments are lower than meta-evaluations.
- Among all of 1,596 checklist items, items concluded as “meta-evaluation > self-assessment” are 146 evaluations (9.1% of population) while items concluded as “meta-evaluation < self-assessment” are 163 evaluations (10.2%).
- Meta-evaluation judged the quality of internal ex-post evaluation relatively lower than that of self-assessment in general.
- Effectiveness/Impact and Sustainability (especially in financial aspect) show lower scores of meta-evaluations than self-assessments.
- There might be the difference of understanding of the check items for which either self-assessor or meta-evaluator counts scores nevertheless other evaluator/assessor judged “NA” (e.g. TC’s and GA’s No.4, 19, 25, 26, TC’s No.13 and GA’s No.12,14) .

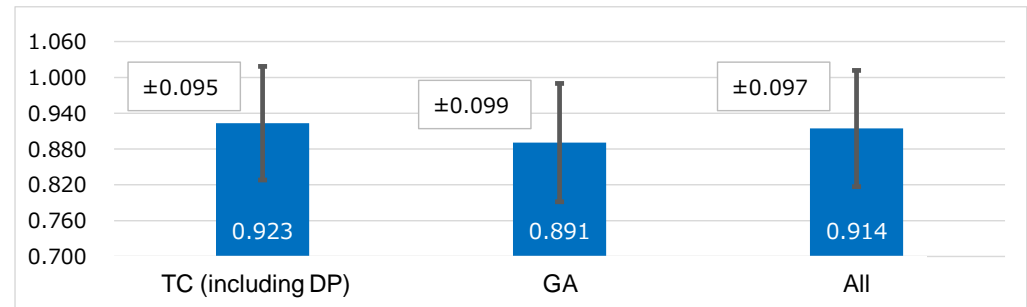


2. Analysis of evaluation criteria

2. Analysis of evaluation criteria (Relevance)

(Relevance) Average, dispersion and standard deviation of 59 targeted projects (standardized score)

	TC	GA	All
Average	0.923	0.891	0.914
Dispersion	0.009	0.010	0.009
Standard deviation	0.095	0.099	0.097



Note : Bar graph shows average of standardized score and error bar (error range) shows standard deviation.

✓ Most of the descriptions do not include major problems.

2. Analysis of checklist items (Relevance)

Checklist Item (TC)	Checklist Item (GA)	Average (raw score)	Comments and analysis by meta-evaluator
(No. 1) [Development policy] Does the evaluation report mention the names of the development policies of the country, the years covered by these policies, and relevant objectives, etc., at the times of both ex-ante evaluation and project completion?	(No. 1) [Development policy] Does the evaluation report mention the names of the development policies of the country, the years covered by these policies, and relevant objectives, etc., at the times of both ex-ante evaluation and ex-post evaluation?	1.88	[comment] ● Concreteness of descriptions about the relevant objectives is insufficient.
(No. 2) [Development needs] Does the evaluation report mention the development needs of the country, the information to indicate the existence of the needs (e.g., description of the context of the project, etc.) at the times of both ex-ante evaluation and project completion?	(No. 2) [Development needs] Does the evaluation report mention the development needs of the country, the information to indicate the existence of the needs (e.g., description of the context of the project, etc.) at the times of both ex-ante evaluation and ex-post evaluation?	1.90	[comment] ● Concreteness of descriptions about the information to indicate the existence of the needs is insufficient.
(No. 3) [Japan's ODA Policy] Does the evaluation report mention the names of the ODA policies for the country, the years covered by these policies, and relevant objectives, etc., at the time of ex-ante evaluation?	(No.3) Same as left	1.63	[comment] ● Concreteness of descriptions about the relevant objectives is insufficient.

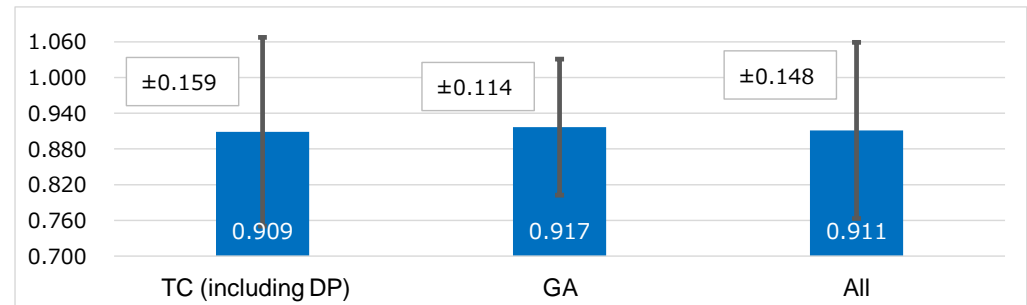
2. Analysis of checklist items (Relevance)

Checklist Item (TC)	Checklist Item (GA)	Average (raw score)	Comments and analysis by meta-evaluator
(No. 4) [Appropriateness] In case the project experienced a problem for Effectiveness/Impact and/or Sustainability, does the evaluation report mention the appropriateness of the project?	(No.4) Same as left	0.60	<p>[comment]</p> <ul style="list-style-type: none"> Some of the evaluation reports do not mention the appropriateness of the project in case the project experienced a problem for Effectiveness/Impact and/or Sustainability. <p>[analysis]</p> <ul style="list-style-type: none"> Most of evaluator tick Not Applicable.
(No. 5) [Evaluation Result] Is the evaluation judgment on Relevance as a whole consistent with the judgments on (1) the development policy, (2) the development needs and (3) Japan's ODA policy (and, if applicable, (4) appropriateness of the approach), at the time of ex-ante evaluation?	(No.5) Same as left	2.00	<p>[analysis]</p> <ul style="list-style-type: none"> There are little descriptions with any problems.

2. Analysis of evaluation criteria (Effectiveness/ Impact)

(Effectiveness/ Impact) Average, dispersion and standard deviation of 59 targeted projects (standardized score)

	TC	GA	All
Average	0.909	0.917	0.911
Dispersion	0.025	0.013	0.022
Standard deviation	0.159	0.114	0.148



Note : Bar graph shows average of standardized score and error bar (error range) shows standard deviation.

✓ Among all criteria, Effectiveness/Impact has the largest variation of description volume, detail and quality.

2. Analysis of checklist items (Effectiveness/ Impact)

Checklist Item (TC)	Checklist Item (GA)	Average (raw score)	Comments and analysis by meta-evaluator
<p>(No. 6) [Project Purpose] Does the evaluation report present and properly analyze the indicator data through the time of project completion?</p>	<p>(No. 6) [Effectiveness] With respect to the project objective(s), does the evaluation report present and analyze the indicator data for the period between the project completion and the ex-post evaluation (including the target year)? If the report fails to include certain indicator data, does it offer an analysis based on supplemental information designated in the EFM? Does the report present the information on qualitative effects (including effects of the soft component if applicable)?</p>	<p>1.78</p>	<p>[comment]</p> <ul style="list-style-type: none"> ● Judgments of project purpose are not consistent with results of achievement of quantitative indicators ● Grounds of judgments of qualitative indicators' achievement are not clear. ● The effectiveness of soft components of GA are not mentioned. ● Indicators had not been set at the beginning of the projects. <p>[analysis]</p> <ul style="list-style-type: none"> ● In case that the target of the indicator is not set, ex-post evaluator has to judge the effectiveness/impact based on the supplemental information such as alternative indicators those are determined by ex-post evaluator in reference to basic examples of indicators and major lessons learned (instructed in JICA's manual), also on the findings in the ex-post evaluation process and changes due to the project, and others. However, it is difficult for meta-evaluator to judge whether the analyzing process is objective and fair. ● It is difficult to make convincing evaluation judgment for ex-post evaluator in case that the initial indicators are not appropriate, vague or abstract. In those cases, the scores of meta-evaluation tend to be lower. ● There are the cases that collection of data is insufficient or supplemental information is incomplete. However, regarding the some of these cases, it also seems practically difficult to collect sufficient data.

2. Analysis of checklist items (Effectiveness/ Impact)

Checklist Item (TC)	Checklist Item (GA)	Average (raw score)	Comments and analysis by meta-evaluator
(No. 7) [Continuation Status of Project Effects] Does the evaluation report present and properly analyze the data showing continued effects (Project Purpose and, if necessary, Outputs)? If the report fails to include certain indicator data, does it offer an analysis based on supplemental information designated in the EFM?	(No. 7) [Effectiveness] With respect to the operation status of the facilities/equipment developed under the project through the time of ex-post evaluation, does the evaluation report present and analyze quantitative data such as capacity utilization, etc. or qualitative information?	1.85	<p>[analysis] 【TC】</p> <ul style="list-style-type: none"> ● Presentation of indicator data is not always necessary. For examples, in order to assess the continuous effectiveness of the training, students' technical level, utilization of knowledge and other information will be able to indicate the effect rather than just the number of students even if the number is set as indicator. <p>【GA】</p> <ul style="list-style-type: none"> ● Some reports mention the operation status of the facilities/equipment at the part of Sustainability.
(No. 8) [Overall Goal] Does the evaluation report present and correctly analyze the indicator data through the time of ex-post evaluation? If the report fails to include certain indicator data, does it offer an analysis based on supplemental information designated in the EFM?	(No. 8) [Impact] With respect to the expected impact (the part after "thereby contributing to "in" Objectives of the Project"), (1) does the evaluation report present and analyze quantitative data or qualitative information through the time of ex-post evaluation?	1.78	<p>[analysis]</p> <ul style="list-style-type: none"> ● Some reports do not mention clearly that the achievement of overall goal is led by the outcome of the project. It seems that, some of the achievements are results of external factors.

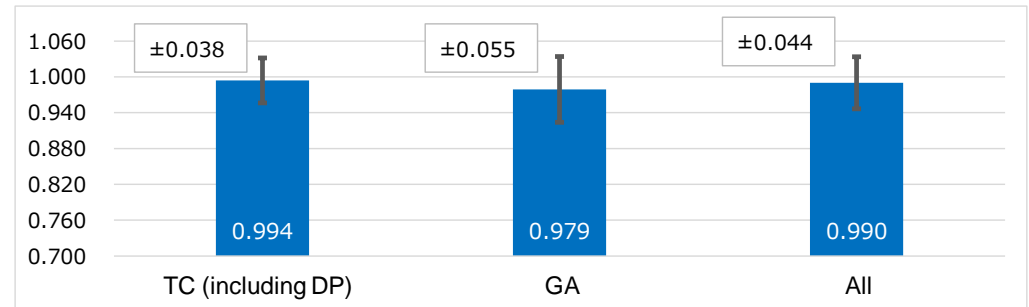
2. Analysis of checklist items (Effectiveness/ Impact)

Checklist Item (TC)	Checklist Item (GA)	Average (raw score)	Comments and analysis by meta-evaluator
(No. 9) Are the grounds for judging individual sub-items of Evaluation/Impact objective and neutral (i.e., Is the judgment not made only based on an opinion of one person or impression? Is the objective not judged as achieved without reasonable explanations even though the indicator values substantially fall short of target values? Is the objective not judged as achieved even though the target values or actual values are unavailable? No it isn't => tick "Yes", Yes it is => tick "No")	(No. 9) Same as left	1.86	[analysis] <ul style="list-style-type: none"> The determination of sample size of sample survey seems not always to be appropriate. It seems to be inappropriate to determine the effect of whole project based on the sample survey when the number of recipients and respondents are much smaller than the number of beneficiaries.
(No. 10) Does the evaluation report state reasons for achievement/non-achievement on individual sub-items and for the increase/decrease in indicator values?	(No. 10) Same as left	1.85	[comment] <ul style="list-style-type: none"> Some reports do not mention the reasons. [analysis] <ul style="list-style-type: none"> Many reports indicate the reasons for non-achievement but do not state reasons for achievement.
(No. 11) [Evaluation Result] Is the evaluation judgment on Effectiveness/Impact as a whole consistent with the judgments on (1) the degree to which Project Purpose has been achieved. (2) the status of the continued effects of Project Purpose toward achieving Overall Goal, and (3) the degree to which Overall Goal has been achieved?	(No. 11) [Evaluation Result] Is the evaluation judgment on Effectiveness/Impact as a whole consistent with the judgments on (1) the quantitative and qualitative effects(Effectiveness) and (2) the expected impact? (3)If other positive/negative impacts are stated and they are considered significant, does the evaluation judgment reflect them?	1.81	[analysis] <ul style="list-style-type: none"> Most of reports evaluate the effectiveness based on the achievement of indicators. However, there are little cases considering the logic between output, outcome and impact, and examined the process of the realization of the project's effect.

2. Analysis of evaluation criteria (Efficiency)

(Efficiency) Average, dispersion and standard deviation of 59 targeted projects (standardized score)

	TC	GA	All
Average	0.994	0.979	0.990
Dispersion	0.001	0.003	0.002
Standard deviation	0.038	0.055	0.044



Note : Bar graph shows average of standardized score and error bar (error range) shows standard deviation.

- ✓ The average score is relatively high while the variation is relatively small. Most of the descriptions do not include major problems.

2. Analysis of checklist items (Efficiency)

Checklist Item (TC)	Checklist Item (GA)	Average (raw score)	Comments and analysis by meta-evaluator
-	(No.12) In case there is a significant gap between the planned and actual outputs, does the evaluation report mention it along with reason(s) in "Efficiency" or other sections (such as "Project Outline" and "Effectiveness/Impacts") ?	1.60	[comment] <ul style="list-style-type: none"> ● In some reports, explanation of reason is insufficient even in case that there is a significant gap between the planned and actual outputs.
(No.12) In case there was a significant discrepancy between the planned and actual inputs (project period and project cost), does the evaluation report mention the reason(s)?	(No.13) In case there was a significant discrepancy between the planned and actual inputs (project period and project cost), does the evaluation report mention the reason(s)?	1.93	[comment] <ul style="list-style-type: none"> ● In some reports, explanation of reason is insufficient in case that there is a significant discrepancy between the planned and actual inputs.

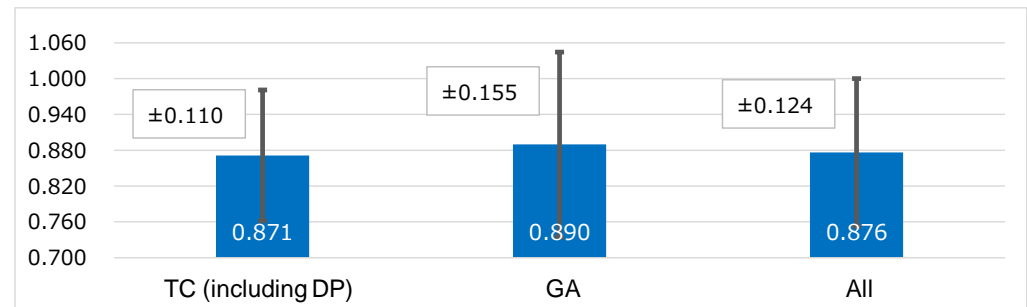
2. Analysis of checklist items (Efficiency)

Checklist Item (TC)	Checklist Item (GA)	Average (raw score)	Comments and analysis by meta-evaluator
(No.13) If the assessment of the inputs is made based on the increase or decrease in the outputs rather than on a simple comparison between the planned and actual inputs, does the evaluation report explain it in relation to the outcomes? Is the explanation reasonable?	(No.14) If the assessment of the inputs is made based on the increase or decrease in the outputs rather than on a simple comparison between the planned and actual inputs, does the evaluation report provide logical explanations including the reasons for choosing that type of assessment?	2.00	[analysis] <ul style="list-style-type: none"> ● There is only 1 evaluation report that is judged based on the increase or decrease in the outputs. ● Most of reports chose basic measure (a simple comparison between the planned and actual inputs). It is because of the difficulty to judge if the increased/decreased outputs are balanced with the inputs.
(No.14) [Evaluation Result] Is the evaluation judgment on Efficiency as a whole consistent with the judgments on (1) the project period and (2) the project cost?	(No.15) Same as left	2.00	[analysis] <ul style="list-style-type: none"> ● There are little description with any problems. (The evaluation judgments are consistent with the judgments on the project period and the project cost.)

2. Analysis of evaluation criteria (Sustainability)

(Sustainability) Average, dispersion and standard deviation of 59 targeted projects (standardized score)

	TC	GA	All
Average	0.871	0.890	0.876
Dispersion	0.012	0.024	0.015
Standard deviation	0.110	0.155	0.124



Note : Bar graph shows average of standardized score and error bar (error range) shows standard deviation.

✓ The average score is low while the variation is large. Among all criteria, Sustainability has the lowest average score.

2. Analysis of checklist items (Sustainability)

Checklist Item (TC)	Checklist Item (GA)	Average (raw score)	Comments and analysis by meta-evaluator
(No.15) [Policy Aspect] Does the evaluation report mention the names of the development policies of the country, the target years of the policies, and relevant objectives, etc., at the time of ex-post evaluation?	-	2.00	[analysis] <ul style="list-style-type: none"> ● There are little descriptions with any problems.
(No.16) [Institutional Aspect] Does the evaluation report mention the organization(s) responsible for sustaining the project effects, and make the roles of such organization(s) in sustaining the project effects clear?	(No. 16) [Institutional Aspect] Does the evaluation report mention the organization(s) responsible for operation and maintenance after project completion, and make the roles of such organization(s) in sustaining the project effects clear?	1.95	[comment] <ul style="list-style-type: none"> ● Most of reports mention the organizations' responsibility and role for sustaining the project effects. ● It seems easy to understand when the reports provide paragraphs for each organization and mention their responsibilities, roles and other information separately.
(No.17) [Institutional Aspect] With respect to the grounds on which the organizational structure and the number of staff are judged to be sufficient/insufficient, does the evaluation report show data in the explanation, and make it clear for what specific activities/services (e.g., dissemination of the model) the reported institutional conditions are sufficient/insufficient?	(No. 17) [Institutional Aspect] Does the evaluation report explain with data the grounds on which it judges the organizational structure and the number of staff to be sufficient/insufficient?	1.76	[comment] <ul style="list-style-type: none"> ● In some reports, the information which support the judgment of institutional sustainability to be sufficient or insufficient are not enough. [analysis] <ul style="list-style-type: none"> ● Most of reports present data such as number of staff and analyze accurately, but there are difference among the reports in the degree of preciseness of data. ● In many cases, ex-post evaluators judge the sufficiency of institution based on the information from interviews.

2. Analysis of checklist items (Sustainability)

Checklist Item (TC)	Checklist Item (GA)	Average (raw score)	Comments and analysis by meta-evaluator
(No.18) [Technical Aspect] With respect to the grounds on which the technical levels are judged to be sufficient/insufficient, does the evaluation report show data in the explanation, and make it clear for what specific activity/service (e.g., dissemination of the model) the reported technical levels are sufficient/insufficient?	(No. 18) [Technical Aspect] Does the evaluation report explain with data the grounds on which it judges the technical level to be sufficient/insufficient?	1.88	<p>[comment]</p> <ul style="list-style-type: none"> ● In some reports, the information which support the judgment of technical sustainability to be sufficient or insufficient are not enough.. <p>[analysis]</p> <ul style="list-style-type: none"> ● In many cases, ex-post evaluators judge the sufficiency of technical level based on the fact if the counterpart organization conducts trainings, prepare manuals, and also the continuity of project activities as well as the information from interviews.
(No.19) [Financial Aspect] With respect to the grounds on which the financial condition is judged to be sufficient/insufficient, does the evaluation report show data in the explanation? If the report fails to include and analyze the budget data, is the financial condition not judged to be "no problem" without presenting supplemental information? (No it isn't => tick "Yes"; Yes it is => tick "No") (4) Does the report make it clear for what specific activity/service (e.g., dissemination of the model) the reported financial condition is sufficient/ insufficient?	(No. 19) [Financial Aspect] Does the evaluation report explain with data the grounds on which it judges the financial condition to be sufficient/insufficient? If the report fails to present and analyze the budget data, is the financial condition not judged as "no problem," etc., without presenting supplemental information? (No it isn't => tick "Yes"; Yes it is => tick "No")	1.51	<p>[comment]</p> <ul style="list-style-type: none"> ● Budget data information are not obtained in some cases. ● Analysis of alternative information are not implemented enough to cover insufficient data. ● Budget data information are obtained but not precisely enough to judge if the financial sustainability is sufficient or insufficient. <p>[analysis]</p> <ul style="list-style-type: none"> ● The average raw score is relatively low compared to other aspects. ● Manual requires ex-post evaluator to analyze sufficiency of budget based on the comparison between budget and actual for past several years. However many reports judge only from budget data.

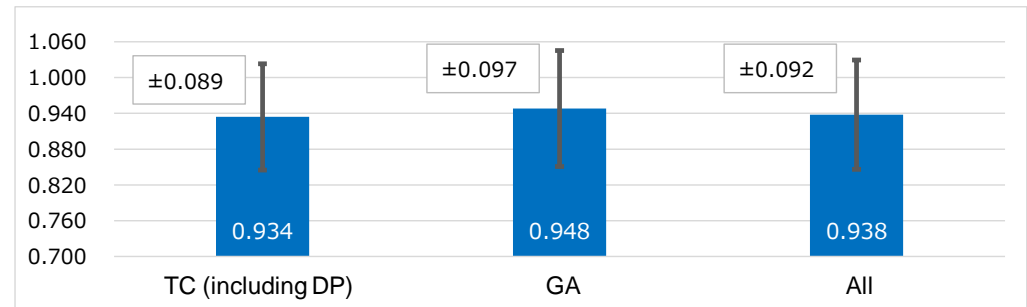
2. Analysis of checklist items (Sustainability)

Checklist Item (TC)	Checklist Item (GA)	Average (raw score)	Comments and analysis by meta-evaluator
(No.20) If issues are found in individual aspects, does the evaluation report mention the factors behind such issues and prospects?	(No. 20) Same as left	1.35	[comment] <ul style="list-style-type: none"> ● The evaluation report does not mention the factors behind issues and also the prospects.
(No.21) [Evaluation Result] Is the evaluation judgment on Sustainability as a whole consistent with the judgments on (1) policy aspect, (2) institutional aspect, (3) technical aspect and (4) financial aspect?	(No. 21) [Evaluation Result] Is the evaluation judgment on Sustainability as a whole consistent with the judgments on (1)institutional aspect, (2)technical aspect and (3)financial aspect?	1.92	[analysis] <ul style="list-style-type: none"> ● Most of the descriptions do not include major problems.

2. Analysis of evaluation criteria (Conclusion/ Recommendations/ Lessons Learned)

(Conclusion/ Recommendations/ Lessons Learned) Average, dispersion and standard deviation of 59 targeted projects (standardized score)

	TC	GA	All
Average	0.934	0.948	0.938
Dispersion	0.008	0.009	0.008
Standard deviation	0.089	0.097	0.092



Note : Bar graph shows average of standardized score and error bar (error range) shows standard deviation.

- ✓ The average score become relatively high. The scores of meta-evaluation tend to be high, because there are many checklist items those are difficult to be judged objectively and fairly as meta-evaluator.

2. Analysis of checklist items

(Conclusion/ Recommendations/ Lessons Learned)

Checklist Item (TC)	Checklist Item (GA)	Average (raw score)	Comments and analysis by meta-evaluator
(No. 22) [Summary of Evaluation] Is the overall evaluation judgment consistent with the description of evaluation results?	(No. 22) Same as left	2.00	[analysis] <ul style="list-style-type: none"> ● There are little descriptions with any problems.
(No. 23) [Recommendations] Does the evaluation report mention in relevant sections the reason(s) why the proposed recommendation should be followed (i.e., issues, etc.)? Is the recommendation concrete enough for the implementing agency to actually implement?	(No. 23) Same as left	1.69	[comment] <ul style="list-style-type: none"> ● The relevance between recommendations and other information stated in the report is not always clear.
(No. 24) [Lessons Learned] Does the evaluation report mention in relevant sections the finding from which you drew the lessons learned? Are the lessons learned both concrete (i.e., indicating what to do in what condition) and general (i.e., being applicable to other projects) so that it can serve as a reference for similar projects? In other words will it be useful if you are to plan and implement a similar project?	(No. 24) Same as left	1.91	-

Third-party Q/C of individual projects

2. Analysis of checklist items

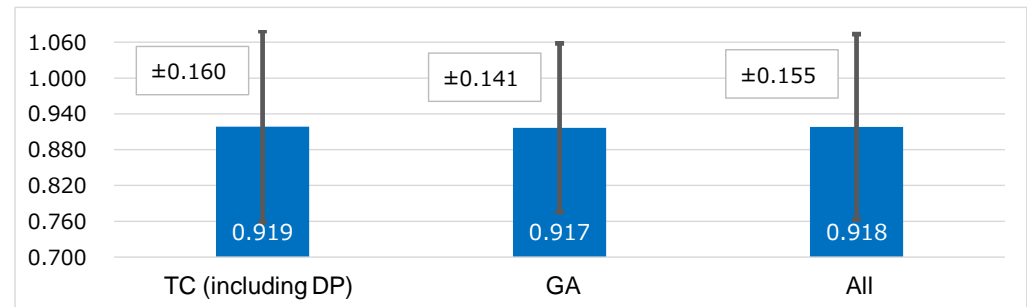
(Conclusion/ Recommendations/ Lessons Learned)

Checklist Item (TC)	Checklist Item (GA)	Average (raw score)	Comments and analysis by meta-evaluator
(No. 25) In case of no description of recommendation or lessons learned, is such way of description adequate? (Example of inadequate description: There is significant problem in impact or sustainability, but there is no recommendation for coping with the problem; There is factor analysis that might lead recommendation, but there is no recommendation derived, etc.)	(No. 25) Same as left	0.78	[analysis] <ul style="list-style-type: none">● Most of reports are ticked "NA" on this checklist item by meta-evaluation, because there are just a few reports without recommendation / lessons learned. Thus, the cases those are ticked "No" on this item much influence on the average score of meta-evaluation.

2. Analysis of evaluation criteria (Overall description)

(Overall description) Average, dispersion and standard deviation of 59 targeted projects (standardized score)

	TC	GA	All
Average	0.919	0.917	0.918
Dispersion	0.026	0.020	0.024
Standard deviation	0.160	0.141	0.155



Note : Bar graph shows average of standardized score and error bar (error range) shows standard deviation.

- ✓ The variation is relatively large compared to other criteria. The influence of the cases that the point was deducted by meta-evaluation becomes relatively larger than other criteria, and the variation tends to be large, because number of checklist items is small and many of them are ticked on "Not Applicable" by meta-evaluation.

2. Analysis of checklist items (Overall description)

Checklist Item (TC)	Checklist Item (GA)	Average (raw score)	Comments and analysis by meta-evaluator
(No. 26) If there are important constraints on evaluation (e.g., inappropriate indicator(s), inadequate information collection, etc.), does the evaluation report explain them and how they might have affected evaluation judgment?	(No. 26) Same as left	1.98	[analysis] <ul style="list-style-type: none"> The detail of the descriptions of constraints and special perspectives are different depending on each ex-post evaluators.
(No. 27) Does the evaluation report identify the sources for all information it presents? ("According to the implementing agency" is acceptable)	(No. 27) Same as left	1.85	[comment] <ul style="list-style-type: none"> Sources of information are described in one gathered part (under the table), thus it is not clear from which source each data comes respectively.
(No. 28) In case the evaluation report states it conducted a sample survey, does it mention the number of valid responses and attributes of the respondents?	(No.28) Same as left	1.22	[analysis] <ul style="list-style-type: none"> There are few cases ticked on this item. Thus, the cases that descriptions about attributes of the respondents are insufficient, much influence on the low average of raw score. In many cases, it is difficult for meta-evaluator to judge if the type of research is basic sample survey or another methodology such as key informant interview.

This report compiles the information collected through the study entrusted by Japan International Cooperation Agency. The information contained herein is not necessarily intended to address the circumstances of any particular individual or entity. Although we endeavor to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. This report was submitted only to Japan International Cooperation Agency. KPMG AZSA LLC will not bear any direct nor indirect responsibility for the use of this report by any third party who accessed this or copies thereof.