

SC assessment and comments on the PM Results indicators

2006 data

11 May, 2007

OUTPUTS

Background

The Output indicator is based on self-reporting of the achievement of output targets that are defined as: *the annual deliverables, defined by quantity and type, expected in a specific year and contributing to achieving the MTP Project Outputs. Output targets are deliverables in the following categories: materials, policy strategies, practices, capacity, and other kinds of knowledge.*

For the first time the 2006 output targets reported in 2007 were based on the MTPs made for 2006-2008 that were designed to link with the PM system.

The Output indicator is a self-assessment indicator and the reported achievement of output targets is audited by an external auditor. The Centers were instructed to report the achievement of output targets listed in MTP 2006-2008 for 2006. Only full achievement was to be reported, not partial achievement. Centers were further instructed to report on output targets that describe deliverable research and capacity building achievements. In previous MTP and PM feed-back the SC has elaborated the types of targets that are not considered acceptable.

The Centers were also given an opportunity to provide an explanation a) for the output target not being achieved due to exceptional circumstances; and b) new output targets.

Process

In the PM system, the SC is invited to provide a short commentary on each Center's Output submission and results.

The SC assessed the output targets and determined the number of unacceptable output targets for each Center. In the SC's commentary the PM indicator for achieved outputs is based on the acceptable output targets only.

SC considered as unacceptable output targets: outcomes, process milestones, on-going activity, vague partnership arrangements, general support/backstopping and technical assistance to partners, and routine genebank management. Following communication with Centers about the visibility of ex post impact assessment studies, the SC included such studies among the acceptable output targets, but encourages them to be reported within indicator 3A. Workshops were accepted as capacity building.

The SC also considered the following explanations provided by the Centers for not achieving the output targets:

- departure of staff member
- delays in inputs
- shortage of seed

- administrative changes in partner country
- time, staff and funding constraints
- merging or subsuming of activity cancellation of course
- government ban
- activity deferred
- project considerably expanded
- slow EU administrative procedures
- cancellation of a joint project to carry out the work
- delays in finalizing bylaws

The SC concluded that only extreme situation in the country of action was an acceptable reasons for not achieving the output targets. In general, the SC felt that accepting the types of explanations given would defeat the purpose of the indicator and would create disincentives to trying to overcome the constraints to achievement, which is the art of good research management. It would be unfair towards Centers that have achieved in spite of facing similar constraints.

Results and observations

The SC observed a much improved quality of the output targets over the earlier years. On average, 89% of the output targets were acceptable. The percentage of achievement of output targets for the Centers, based on either those reported by the Center or those judged acceptable by the SC, are given in Table 1. Calculation of achievement on the basis of acceptable targets did not improve or diminish the result more than 2.9 percentage points.

Table 1 - PM Output indicator results

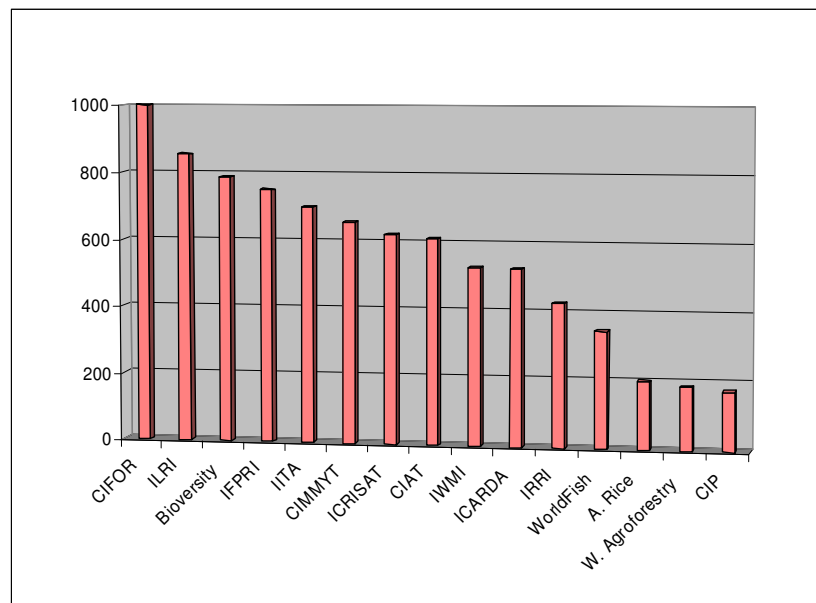
Center	Output targets reported	Not acceptable output targets (#)	Acceptable output targets (%)	Output targets reported as achieved (%)	Achievement of acceptable output targets (%)
A. Rice	71	12	83	94.4	91.5
Bioversity	66	19	71	87.9	89.4
CIAT	84	9	89	91.7	90.7
CIFOR	21	4	81	90.5	88.2
CIMMYT	63	9	86	88.9	88.9
CIP	132	7	95	95.5	95.2
ICARDA	56	5	91	91.1	90.2
ICRISAT	57	9	84	96.5	95.8
IFPRI	64	11	83	90.6	90.6
IITA	75	4	95	89.3	91.5
ILRI	48	1	98	83.3	83.0
IRRI	91	4	96	80.2	81.2
IWMI	51	2	96	90.2	89.8
W. Agroforestry	176	23	87	79.6	80.4
WorldFish*	63	5	92	73.0	75.4

* One unachieved output targets was excluded because of legitimate reason given by the Center for non-achievement

On average, the Centers reported very high level achievement of output targets. As noted earlier there has been a marked improvement in the planning and definition of MTP outputs and in achieving them in the planned time frame. The SC views this improvement also as an indication of a better understanding of the impact pathway. The Centers, through the MTPs, have become more results oriented and the planning for results more transparent.

The SC notes that while all Centers have a reasonably high level of achievement there remain differences among the Centers in the resources required per output target (see Figure 1), with the highest being more than five times that of the lowest. The SC notes that there could be a hazard in this self assessed indicator to set easily achievable targets in order to then ensure a high percentage of achievement. Also there is a possibility that Centers differ in the “scientific risk” undertaken to address the more intractable problems. The SC commentary on MTPs 2006-2008 did not include an assessment of the magnitude of output targets or on the ambition of the plans. The challenge for this indicator is to maintain a healthy level of ambition in research planning and acknowledge serendipity in research.

Figure 1 – Average cost per output target (US\$ '000)



The SC notes that the MTPs provided by the SWEPS often plan for process targets (organising meetings, better communication, facilitation of the steering committee work etc.). The SC may need to reconsider whether this PM indicator is appropriate for the programs that are intended for facilitation and coordination, and how such roles are best reflected in the MTPs.

Short commentaries providing a brief assessment of output target quality and an accordingly adjusted indicator value were inserted in the PM system for Outputs.

OUTCOMES

Background

In the CGIAR MTPs and PM system, Outcome is defined as *the external use, adoption, or influence of a Center output(s) (e.g. by partners, stakeholders, clients).*

Centers were instructed to report on the **five (5) most significant outcomes in 2006** resulting from outputs that the Center produced in 2003-2005. Each reported outcome needed to be supported by evidence of credible documentation from either the Center or target user. In addition to the outcome statement, Centers were requested to provide the following information in each outcome description:

- What output(s) resulted in the outcome and in which MTP was the output first identified?
- Who used/adopted or was influenced by the output? Identify the kinds of people or entities affected and their location by country.
- How was the output used or adopted? What was the nature of the influence of the output?
- What is the evidence for this outcome? Specifically, what kind of data were collected or study was conducted? Who collected the data and/or conducted the study?

In 2007, the SC assessed the outcome statements using two criteria.

- Criterion 1: Nature of the reported outcome and validity of the verifiable source of evidence. Each case was scored 1 or 0 based on an acceptable outcome and a verifiable evidence. Total possible score for this item = 5.¹
- Criterion 2: Clarity of describing the outcome. Each case was scored 1 or 0 based on the clarity and specificity. Total possible score for this item = 5.

Note: the cases that scored 0 on Criterion 1 also scored 0 on Criterion 2.

This provides for a maximum score of 10 for the PM indicator for outcomes.

In the course of assessment the SC observed a number of border-line cases. Thus the SC also used the following to moderate their assessment in these cases:

- What if the research results are older than from 2003-2005 as requested in the instructions? *If it is likely that the outputs are sufficiently recent the lack of explicit date of the output does not matter this year.*
- What if the evidence is not directly relevant, but the statement is completely clear? *The statement can be given full points depending on the specificity of the description.*
- What if neither the statement nor the evidence fully link the outputs to the outcomes (although the linkage may be likely)? *The Center can get 1 point maximum depending on the specificity of the statement.*
- What if the output has been mainly created elsewhere, and the Center has simply modified it before delivering it for use. *If there is a significant modification relevant for adoption of the*

¹ Note. This was a change from the previous year where availability of verifiable evidence was a third criterion and subsequently the maximum total score was 15

output the statement can get 1 or 2 points depending on the specificity of the statement and evidence. If the Center is only a distributor, the statement gets 0 points.

- What if the outcome is a very early, pilot outcome - almost like feed-back during the research process? *This is part of the process leading to outcome, but not an outcome.*
- What if the outcome is a small example of many possible outcomes from same research? Can Centers continuously keep reporting different outcomes from same research? *In general, Centers should not report incremental additional outcome from the same output in subsequent years. Depending on the nature of the output, the statement could get points.*
- What if the outcome results from research results accumulated over a long time and it is unclear whether the latest modification resulted in the outcomes reported? *The score is reduced if the linkage between the output and the outcome is not clear.*
- What if the Center either makes reference to evidence that is not immediately available or sends it to the SC afterwards? *The SC attempts to obtain relevant evidence from the Center, but does not either screen generic sources or contact individuals in search of suitable evidence.*

The SC will consider these observations in revising the Outcome instructions for 2008.

Results

The scores for the PM indicator for outcomes for the Centers are given in Table 2. Comments on individual Center outcome statements are presented in Table 3 at the end of this section. The scores ranged from 2 (one Center) to the maximum 10 (two Centers). The mean score for all Centers is 7.6 and the median is 7.

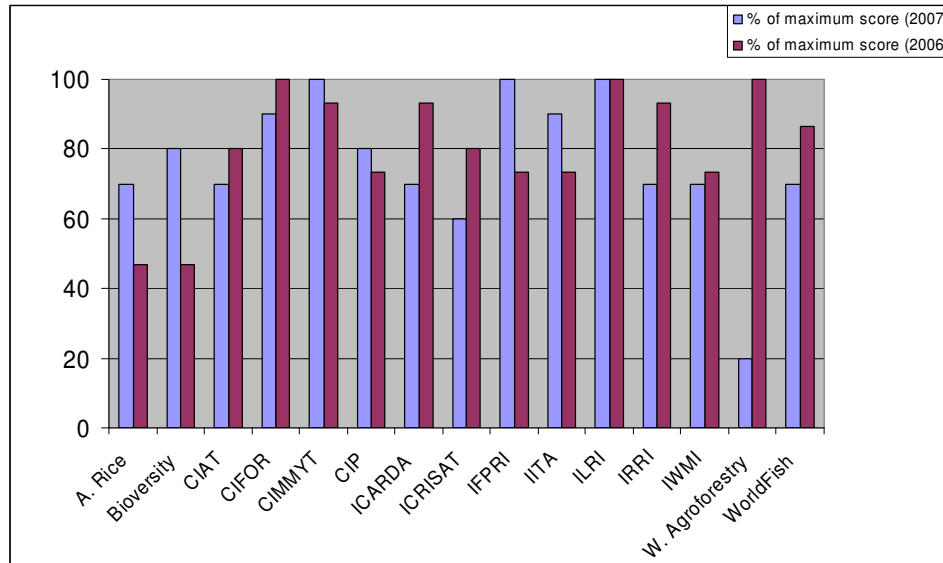
Table 2 – Outcome scores

Center	Score
Africa Rice Center	7
Bioversity	8
CIAT	7
CIFOR	9
CIMMYT	10
CIP	8
ICARDA	7
ICRISAT	6
IFPRI	10
IITA	9
ILRI	10
IRRI	7
IWMI	7
World Agroforestry Center	2
WorldFish Center	7

There is a slight reduction in the overall results for the outcome indicator this year compared to last (Figure 2). The average score in 2007 was 76% of the maximum score compared to 80.9% last year. The scores of this year, however, are not directly comparable with those of 2006 due to the change in the rating scale. In both years 66% of all outcome statements received the full score. However there is an increase in the number of poorly rated outcomes

(14.7% in 2007 were score 0 compared to 9 % in 2006). This may have been due to the stricter requirement to report outcomes from research results received in a defined period.

Figure 2 – Comparison of outcome results in 2007 and 2006



The reasons for a low score in 2007 included:

- Statement describes process, not external use, influence or adoption of Center outputs
- Weak (or non existent) link between Center outputs and the outcome
- Outcome from old research activities
- Very preliminary outcome from on-going research or other activity
- Statement is of potential outcome only

Many of the evidence documents were not directly relevant for the specific Center output leading to the described outcome. Much of the evidence was by association, including reports on different events with only a weak association to the outcome described, receipts, photos, personal contact information, newspaper articles, and generic sources of evidence, such as NARS reports, NGO reports, donor report, FAO statistics, Radio and press reports. Centers are encouraged to gain more knowledge on the plausible impact pathways including the documentation of possible outcomes in order to improve the monitoring and uptake of their research.

The SC notes that many statements describe outcomes from “local” goods activities. The SC would like to see more focus on outcomes from IPG research. The SC encourages the Centers to articulate the pathways for impact, which would assist them in clarifying who are the likely “users” of their IPG research outputs and what constraints will need to be overcome to have a genuine outcome.

The indicator currently requests the same number of outcome statements from each Center and thus might favour the larger over the smaller Centers. However, there is no indication

that the smaller Centers had more difficulties in identifying 5 good outcome cases annually. Also the indicator does not measure the magnitude of the outcomes. As stated earlier the assessment criteria focus on the output-outcome linkage and the specificity of the statement about the outcome. The SC will look at the feasibility in adding a dimension of magnitude of the outcome for future assessment.

Table 3 - Assessment of individual Center output statements

Outcome statement	Assessment/comments
Africa Rice	
1 – Rice policy change in Nigeria	Score =0. The statement is not clear about the developments in Nigeria regarding policy having resulted from a WARDA intervention, nor does the evidence indicate that. Referenced evidence sources are generic and not possible to check; e.g. donor operational reports.
2 – A WARDA tool for farmer varietal exposure adopted by NARS	Score=1. Adoption of varieties as result from PVS was reported already in 2006. Rather than a specific output from WARDA PVS seems a strategy used to increase the adoption of new varieties, which is likely to be continuous used. Score of one is given as it is likely that the NARS have adopted the specific PVS method as modified by WARDA. Reporting incremental and equal outcomes on the same output repeatedly is not acceptable (more and more varieties over years adopted using the same PVS method). Generic evidence sources are included; e.g. “NARS reports”.
3 – Information systems providing a template for development	Score=2. The statement makes it clear that NIVISA system was based on WARDA’s product WAIVIS and is being used by participating countries. The evidence document is a proposal from Togo ministry to UNDP for development of a water resource database. The actors, a NARI and a NGO in Togo had been involved in NIVISA, and with that experience had got involve in the UNDP project on dams and reservoirs in Togo. Generic evidence sources are included; e.g. “NARS, NGO reports”.
4 - Partner organizations in Africa use video to underpin scaling-out of rice technologies	Score=2. The stated outcome is use of communication method developed elsewhere and fine-tuned by WARDA and partners for specific West-African situation for rice. The evidence is not quite relevant and describes the fine-tuning rather than the uptake or use. Generic evidence sources are included; e.g. “NARS, NGO, radio, press reports”.
5 – Uptake of Upland NERICA Varieties by NARS in SSA	Score =2. New uplands NERICA varieties have been adopted. The evidence document about uplands NERICAs is from Nigeria where WARDA was responsible for the dissemination of the varieties in a GATSBY funded project. Considerable and rapid adoption of NERICA1 is documented. Generic evidence sources are included; e.g. “FAO documents, press reports”.
Bioversity	
1 – The Nepal Seed Law was modified to allow the release of	Score=2. The statement is specific about national activity and several Bioversity activities. It is very likely that Bioversity

Outcome statement	Assessment/comments
farmer varieties, as a result of recognition of the validity of participatory plant breeding among researchers, farmers, non-governmental organizations and national policy makers in Nepal. The new law was applied to the farmer rice variety Pokhareli Jethobudho	interventions, particularly GRPI combined with national pressure led to the policy change. The evidence document is an assessment of a different global project, with only indirect relevance to this specific case.
2 – A greatly increased amount and diversity of date palm genetic resources, much of it rare or threatened, was conserved on-farm in three countries in the Maghreb (Algeria, Tunisia and Morocco).	Score=2. The statement is very specific about the outcome. Bioversity’s interventions are explained in the evidence report by GEF, which is the evaluation of the project and describes its outcomes.
3 – The adoption of crop and multi-crop descriptors resulted in increased efficiency and effectiveness of <i>ex situ</i> germplasm management.	Score=2. The output (multi-crop descriptors) is clear and its use was comprehensively documented. The evidence is a broader external study that covers also crop descriptor use.
4 – Government partners in Armenia, Bolivia, Madagascar, Sri Lanka, and Uzbekistan are using the IUCN Red List criteria to review the national and global threat status of crop wild relative species.	Score=2. The training activity that led to use of the IUCN Red List criteria is sufficiently elaborated. The evidence which is a project document states the progress following training
5 – A strategic framework for a European Genebank Integrated System (AEGIS) was adopted by the Steering Committee of the European Cooperation Programme for Plant Genetic Resources (ECPGR), which includes 38 European countries.	Score=0. The statement is not about outcome from an output, but change of status of a Bioversity managed project (AEGIS) by a Bioversity coordinated program (ECPGR). This can be considered a process outcome, which is not intended here.
CIAT	
1 – Linking farmers to bean markets in Ethiopia	Score=2. The output, CIAT’s intervention on seed systems has led to the outcome of the new seed systems emerging and bean varieties being adopted and grown. The evidence is a specifically about Ethiopia and the evidence is by association to the case reported. The timing of evidence documents should be more accurate – the outcome is expected to have occurred (documented) in 2006.
2 – Massive, <i>in vitro</i> Propagation Systems of Cassava Adapted to the Needs of Small Farmers in Colombia	Score=2. The statement is specific about the propagation technology and its adoption. The outcome is early, but shows clearly used of an established technology. The evidence is about the technology rather than about the outcome.
3 – A superior <i>Brachiaria</i> hybrid combining drought tolerance,	Score=2. The statement is very clear about release and adoption of CIAT produced <i>Bracchiaria</i> cultivar. The evidence document

Outcome statement	Assessment/comments
resistance to spittlebug and adaptation to acid infertile soils released by a commercial seed company in LAC countries	is about testing <i>Brachiarias</i> in Thailand and not very specific about the described outcome from CIAT outputs.
4 – Developing the multipurpose soybean value chain to increase income of smallholder farm families and other rural entrepreneurs in East Africa	Score=0. The statement describes soybean value chain, but CIAT's intervention and when it may have taken place is not clear at all. The evidence is not relevant for the case but describes testing the effectiveness of farmer participatory breeding of soybeans.
5 – Building Capacity in Gender Analysis and Gender Mainstreaming in the NARS of ASARECA	Score=1. The outputs have been a large number of methodological ones, and outcomes (capacity through gender analysis activities) are not clear from most of them. Outcomes, like adoption of the new capacity into action, are just accruing and only a few are specified. The project evaluation report is quite weak and generic and mostly about project results.
CIFOR	
1 – CIFOR research influenced UNFCCC Subsidiary Body for Scientific and Technological Advice on reducing emissions from deforestation in developing countries	Score=1. The case is more about CIFOR providing multiple types of results and analyses. But the outcome, i.e. the use of these results is not clear. However, it is quite clear that CIFOR is having influence – it is quoted some 30 times in the evidence report.
2 – New regulations adopted to better facilitate conservation and development in Ethiopia	Score=2. The outcome is a result of research and policy dialogue and the linkage from CIFOR intervention to the outcome is clear. The evidence available is an internal media story, but has also been documented in a peer reviewed publication. The steps toward adjustments in forestry legislation are clearly described.
3 – Restrictive regulations on the transport of forest products abolished	Score=2. The statement is very clear about CIFOR's interventions and their influence on national regulation. The evidence available is an internal media story; the details are clear. The kinds of outputs (papers and policy briefs) that have likely contributed to the outcome were targeted in MTP 2005-7 for 2005.
4 – CIFOR's research to help protect natural forests and secure rural livelihoods by improving financial due diligence and impact assessments for forest-related investment	Score=2. The statement is clear. In the evidence CIFOR is sited as having provided the results that have led to the funding decisions and collective action by CSOs.
5 – CIFOR's research helps to target the World Bank's Indonesia Forest Strategy	Score=2. The statement is clear and the evidence supports it.
CIMMYT	
1 – Improved disease resistance, yield potential and diversity by using CIMMYT re-synthesized wheat germplasm	Score=2. Statement describes two steps: Release of an earlier synthetic is being adopted and novel germplasm from defined output is being incorporated. The evidence, which is about survey results on the use of the germplasm (early outcome), is not specific to China which is elaborated in the statement.
2 – Adaptation of rainfed	Score=2. The statement is clear about the adoption of

Outcome statement	Assessment/comments
conservation agriculture models by NARS in Kazakhstan	conservation agriculture practices developed by CIMMYT in Kazakhstan and the evidence supports it.
3 – Laying the foundations for quality wheat in China	Score=2. The statement is clear about CIMMYT’s work on marker development although less so about CIMMYT’s role in the cultivar development, which is however, very likely. The evidence is about the release of the new variety.
4 –Scaling up stress tolerant maize lines in East and Central Africa (ECA)	Score=2. A clear statement about CIMMYT’s new stress tolerant maize released in varieties. Evidence is very relevant: the project review with assessment of outcome and future impacts.
5 – Reaching local partners with improved maize and wheat lines in Saraguro, southern Ecuador	Score=2. The statement is clear about use of CIMMYT germplasm in Ethiopia. The evidence identifies CIMMYT as source of wheat germplasm but it is not clear about maize.
CIP	
1 –Developing new products from native potatoes	Score=2. The kind of knowledge package CIP delivered is not very well elaborated, nor well identified in the MTP. However, the statement is specific about the outcome and the evidence document identifies CIP as the source of the Participatory Production Chain Approach.
2 – Remote sensing techniques and data used in free trade negotiations by Panama	Score=2. The statement about use of remote sensing technology is very specific and the evidence supports it.
3 – Utilization of advanced populations to identify promising clones in variety selection trials in various countries: case of potato	Score=1. The original output has not been identified. CIP has initiated the breeding scheme and supplied germplasm. The evidence shows early positive results but testing is ongoing.
4 –Impact of orange-fleshed sweetpotato on nutrition in Mozambique	Score=2. The CIP intervention is public awareness and training campaign. The evidence document is a research report from the subsequent project implying impacts from OFSP and not very specific about the CIP campaign. Yet, CIP’s influence on the positive behavioral change is evident.
5 – Introduction of sweetpotato as animal feed into the cropping systems in the dry forests in Northwestern Peru	Score=1. The reference to the stated output is not clear and nor is it described in the MTP referenced. The linkage is merely CIP making cultivars available and county releasing them which is less than the outcome statement claims.
ICARDA	
1 – Pilot-village based seed enterprises developed in Afghanistan	Score=2. The ICARDA intervention is a development action (support to seed systems) and the outcome is seed production by the established pilot enterprises.
2 – NARS take up innovative research approaches to livestock development	Score=2. The genuine outcome – added value from a number of small projects in change of culture, i.e. horizontal networking, is better documented in the evidence than in the statement which does not differentiate between action and the outcomes from that action.
3 – Integrated Resource Management approaches used in Syria, Central Asia	Score=1. The statement describes three outcomes from the same multiple outputs on technologies, institutions and policies. Evidence covers one but it doesn’t show the linkage from ICARDA’s interventions.
4 –Thirty new crop varieties released	Score=2. The statement unnecessarily describes multiple outcomes from different outputs and activity areas but the

Outcome statement	Assessment/comments
	essential and straight forward outcome is from varieties deriving from ICARDA materials being released for which there is evidence.
5 – NARS use new approach to measure water-use efficiency	Score=0. The ICARDA output-outcome linkage is not clear nor documented. It is not obvious that the described intervention is a novel approach and the evidence merely documents the water problem and training.
ICRISAT	
1 – China embraces pigeonpeas	Score=0. The stated output is assessment of potential and opportunities. The specified breeding outputs with reference MTP 2005-2007 are clearly not the ones that have led to cultivation of pigeon pea in China which is obviously resulting from much earlier work as seen in evidence material that is dated much before the limit for outputs (2003-2005). There is no evident link between ICRISAT research and establishing a green vegetable trade for pigeon pea.
2 – Watershed IPG spills over from Asia to Africa	Score=0. The statement is not clear about the ICRISAT research intervention and the outcome is potential at most. Output-outcome linkage is lacking. This is about process of moving experience from one continent to another and so far pilot sites for adaptation and demonstration are being established. The outcomes are not observable as yet.
3 – Doubling soybean productivity in India	Score=2. Output is reported although reference is to a very recent MTP; outcome seems potential but there is evidence that national partner has changed conceptually the approach to soybean systems research through the ICRISAT output. Outcome is the use of the watershed approach interaction with soybean. The other research output (not spelt out) is the model for extrapolation. It seems that there has been a change in the Indian Govt to take up the soybean x watershed approach.
4 – Outscaling micro-dosing and “warrantage” in Mali	Score=2. The statement is specific and the evidence supports it. Synthesis of technology uptake processes led to the outcome.
5 – PPP improves fertilizer use	Score=2. The outcome, uptake of ICRISAT recommendation for national testing and positive adoption results, follows from very recent outputs. This represents early outcome.
IFPRI	
1 – Global Hunger Index to help mobilize political will and promote good policies by ranking countries and illustrating trends	Score=2. The statement is clear about how IFPRI’s Global Hunger Index has influenced national government discussions and international agency advocacy. Relevant evidence is compiled into a single document in all IFPRI’s outcome cases.
2 – Ethiopian market reforms	Score=2. The statement specifies IFPRI’s interventions and the resulting Ethiopian Government action. Evidence in a single source both about the intervention and the influence.
3 – Collaborative master’s program of science in agricultural and applied economics run by AERC	Score=2. The statement and evidence specify IFPRI’s intervention leading to a master’s program being run by an African consortium. This represents capacity outcome.
4 – The Regional Network on HIV/AIDS, Rural Livelihoods and Food Security influences local and	Score=2. The statement describes outcomes (influence on governments, donor and trade organization) from IFPRI-led network. RENEWAL is referenced in some evidence but IFPRI

Outcome statement	Assessment/comments
international actors	may have also directly influenced the outcomes.
5 – IFPRI’s WTO research influenced research community and policy makers	Score=2. The statement is specific about IFPRI’s interventions. Evidence shows how policy makers and research community were influenced.
IITA	
1 – Use of cassava standards in Nigeria, Uganda and other countries	Score=2. The statement and evidence is largely about the dissemination of the IITA guidelines but there is evidence of influence as a consequence.
2 – Reduction of child labor incidents on cocoa farms in Ghana	Score=2. The statement specifies IITA intervention that led to change in child labor practices.
3 – Adoption of nutrient management practices in northern Nigeria	Score=2. The outcome from IITA output is clearly indicated in the statement. The journal article documents farmer use and preference.
4 – Selection and release of improved cassava genotypes and populations in Sierra Leone, Benin and Nigeria	Score=2. Statement and evidence give specific information about IITA cultivars released.
5 – Dissemination of improved cassava planting materials in Southern Sudan	Score=1. Dissemination is not a clear outcome from IITA results. It is not clear which output has led to the dissemination. The first evidence is about availability of planting material in Sudan in 2007, and the second is a survey in 2007 of what farmers grow and how it has done – they are not about a specific outcome. The materials are obviously of IITA origin and for that score 1.
ILRI	
1 – Climate vulnerability and poverty in Africa	Score=2. The statement is specific about how results from ILRI’s targeting research influenced activities and funding by other agencies.
2 – Policy and management options for pastoral lands in East Africa	Score=2. The linkage leading to outcome from ILRI research to participatory community research to prepare maps and data that Kenyan authorities then used is specific. Linkage from ILRI’s work to outscaling to Tanzania is not so clear.
3 – Pro-poor dairy development takes regional hold in Eastern Africa	Score=2. The ILRI outputs added to earlier research on dairy policies. The outcome may involve some of the earlier ones but the output-outcome through ASARECA linkage is shown.
4 – Live vaccine safely deployed in pastoral areas of East Africa	Score=2. The output, the infection and treatment method related to vaccine, is clear and the linkage to the clear outcome due to the method is also specific.
5 – Land-use change impacts and dynamics (LUCID)	Score=2. The statement cites the major findings from the ILRI project evaluation report, which support the outcome case.
IRRI	
1 – Two IRRI bred superior New Plant Type rice lines were recommended by the Philippine National Rice Technical Working Group to the National Seed Industry Council for release as commercial national rice varieties	Score=2. The statement is specific about the cultivars released. Evidence only through web-site search.
2 – NARES in India and the Philippines used IRRI derived	Score=2. The statement and evidence is sufficiently elaborate about DNA markers. There are several outputs related to

Outcome statement	Assessment/comments
DNA markers to introgress the sub1 gene into locally popular varieties and used Swarna-sub1 introgression lines directly in field evaluations	submergence tolerance being used.
3 – NARES in Bangladesh, China, India, Indonesia, Myanmar, Philippines, and Vietnam are using the Site Specific Nutrient Management (SSNM) approach in rice research and extension	Score=2. The statement is specific about the adoption of SSNM developed by IRRI.
4 – New rice cultivars, improved germplasm, and germplasm from the genebank are being used by NARES in their breeding programs to develop high-quality, locally-adapted cultivars.	Score=1. The statement describes generic and on-going use of materials from the genebank through INGER nurseries.
5 – NARES in Asia use a research network, the Consortium for unfavourable Rice Environments (CURE) that links their national systems' strategic research to technology development for increased rice productivity and more stable production resulting in the improved wellbeing of the resource-poor rural households in unfavorable environments.	Score=0. Participation in a network and the research conducted within it is not an outcome.
IWMI	
1 – Water use associations transform canal management in Central Asia	Score=2. The statement is specific about the IWMI interventions and it leading to government adoption of institutional reforms.
2 – Co-management of electricity and groundwater in Gujarat, India	Score=2. The statement and evidence are clear about use of strategy in resource management although the basis of the IWMI recommendation is not quite clear.
3 – RUAF outcome story	Score=1. The output, which is not quite clear, is more like advocacy action which has led to recognition of UPA in municipal planning.
4 – Development strategy for irrigation sector of Sri Lanka 2006-2016	Score=1. There is no doubt that IWMI influenced the new strategy, because it seems to have largely written it. The preceding research output is not clearly described.
5 – Influence of IWMI Tata Water Policy Program (ITP) research	Score=1. IWMI's intervention is a program plan which was accepted within the Indian government budget. The water policy program outputs are not quite clear.
World Agroforestry Center	
1 – Adoption of fodder shrubs by dairy farmers in East Africa	Score=0. There is no recent output and the outcome is likely resulting from activities in 1990s. Reference research reports and adoption study report are also old.
2 – Influencing the strategic	Score= 0. ICRAF's "technological output" is not defined and the

Outcome statement	Assessment/comments
options for forest assistance in Indonesia	evidence document does not show whether ICRAF was influential in the development of forestry sector options strategy for Indonesia.
3 – Wide use of participatory domestication of high-value indigenous trees in the humid tropics of West and Central Africa	Score=0. Outputs of ICRAF’s long-term involvement in tree domestication seem old and there is no linkage between outputs and the claimed outcomes, which seems generic “practice of tree domestication”. Some recent uptake is occurring through joint testing which is still Center activity, not yet outcome.
4 – Widespread uptake of Landcare in the Philippines	Score=0. There is no evidence of ICRAF output other than extension type of role starting mid-1990s. The outputs are old, and references are to Australian work only. Adoption has taken place in 2003-4. An impact study does not show attribution to ICRAF.
5 – Mainstreaming integrated natural resource management perspectives into Ugandan level development processes and extension system (NAADS)	Score=2. Although the specifics about the outputs that led to the outcome are not quite clear, the statement is clear about the AHI-led study having influenced a national implementation strategy on integration of natural resources management.
WorldFish	
1 – Capacity strengthened in Mekong region through BayFish	Score=1. The statement is quite clear about managers in Vietnam and Cambodia having used BayFish but the evidence is very lacking. The activities are not clear outcome.
2 – Importance of Genetic management in fish farming	Score=2. The statement and evidence together show the link from the WF breeding approach output to a change in genetic management of Tilapia in Africa.
3 – Low-impact methods for capture and culture	Score=1. The statement defines the outputs and early outcome seems to have occurred, but the evidence seems irrelevant.
4 – Seed supply systems in Bangladesh	Score=2. WF has through participatory research help established decentralized seed supply production centers in villages in Bangladesh. Adoption of technology has increased following FFS training.
5 – Integrated aquaculture-agriculture	Score=1. The outcome; incorporation of IAA into new project designs is clear but there is just little evidence of recent results; generally the outputs on IAA range over time.

IMPACT

3a. Commitment to Documenting Impacts and Building an Impact Assessment Culture

Background

For the 2007 PM exercise (evaluating Centers’ performance in 2006), SPIA had significantly modified the format in which the information requested on ex-post impact assessment related activities was reported. This was in response to lessons learnt from last year’s exercise—which was characterized by incomplete and ambiguous information that SPIA had to subjectively evaluate—and was supported by the views of both SPIA members and the Centers’ Impact Assessment Focal Points (IAFPs) that the submission and evaluation needed

to be more transparent and simplified. Accordingly, SPIA developed a checklist this year to replace the narrative-style request of last year, but essentially asking for the same information. Before finalizing this checklist format, SPIA received from the Centers' IAFPs several suggestions for modifications which were incorporated into the 2007 PMS request.

For this year's checklist format information was requested under three major criteria that SPIA used to score the 3a impact indicator (weights indicated):

- (I) Ex-post impact assessment (epIA) studies / advancement of epIA methods (70%),
- (II) Building an impact assessment culture at the Center (20%),
- (III) Communication / dissemination and capacity enhancement (10%).
- (IV) Other impact related activity (overall modifier)

The complete checklist with weights assigned to the individual components of each major criteria is appended (Annex I).

After the information submitted by Centers for the 2006 3a impact indicator was reviewed, discussed and initially evaluated, SPIA members spent considerable time and effort in evaluating a number of alternative scoring methods. This exercise was necessary to have a transparent and fair process that took into account the differences in sizes of the Centers and considered each Center in relation to benchmarks of achievement (or yardsticks) rather than relative to other Centers' performance. This allows Centers to monitor their own progress from year to year and de-emphasizes performance only in relation to other Centers' results. Thus, unlike the method used last year, the top performing Center (for any given component) no longer constitutes the basis for which all other Centers are scored. After a careful evaluation of the various alternatives, the standing panel adopted a scoring framework that is described below.

Unlike last year, SPIA is not performing a 'verification' role this year with respect to the data submitted by the Centers, apart from making an assessment as to whether a study submitted by a Center actually qualifies as an epIA². This pertains mainly to Criterion I: If a study did qualify as an epIA, all information about the study submitted by the Center was accepted at face value. (SPIA understands that although no auditing/verifying will be done this year for this indicator, it would be next year).

Scoring method

Criterion I: a combination of the volume of epIA studies, quality of epIA studies and investment in epIA related activity was used to compute a score for this criterion as follows:

- a) # of epIAs published per US\$ million of overall average Center budget (constituting 39% of total weight): thus, Centers submitting one or more valid epIA³ per US\$ 5 m of

² SPIA reviewed whether each submitted study, as described by the Centers themselves, qualified as epIA or not. The studies that were considered non-epIA (farmer preference/ demand-type studies, pilot technology evaluations, etc.) were removed from the scoring exercise. For each remaining study, the three SPIA members and the secretary independently tallied scores on a pre-designed excel spreadsheet. Whenever the scores diverged significantly, it was discussed and consensus reached. Scores were given component-by-component for each study and then summed up across all the studies.

³ An epIA study refers to a published journal article, conference paper, book chapter (but not entire edited book), report or any other publication that has entered the public domain, which is not a revised version of an earlier submission.

budget (a SPIA-defined yardstick) received the highest # of points (i.e., 39); those submitting less than that received a proportionate score.

- b) sum of quality characteristics for all the qualified epIA studies submitted (26% of total weight): thus, from a maximum possible score of 26, which characterizes studies with the best (highest scoring) elements in each of the six IB components, a 'per study' quality of epIA score was calculated for each Center.
- c) expenditure devoted to epIA work in relation to the overall average Center budget (5% of total weight): thus Centers investing 1% or more of their total budget to epIA activity received the highest score (i.e., 5), and those investing less than that, a lower score, in direct proportion.

A composite score for Criterion I was calculated by summing the scores from the individual components (a) – (c) described above.

Criterion II: a combination of (i) SPIA's subjective evaluation of the development and application of epIA information and (ii) assessment of the # of internal IA workshops convened to help assess expected impacts of planned/ongoing research. Numeric evaluations (0 to 5 scoring) were given of narratives describing each of the following:

- a) systematic evaluation of user relevance of Center research outputs produced within the past year(5% weight);
- b) use of empirical epIA findings as a basis for quantitative ex-ante impact projections that contribute to the Center's priority-setting procedures (5% weight)
- c) establishment of baseline studies to provide counterfactuals for future epIAs (5% weight).

This aggregate score was combined with

- d) an objective assessment of # of internal workshops such that Centers with one or more workshops per \$10 million of Center budget (a SPIA-defined yardstick) receive the highest score, i.e., 5, and those less than that receives a proportionate score.

Composite score for Criterion II was calculated by summing the scores from (a) – (d) above.

Criterion III: a combination of objective assessments of the number of communication and capacity related events/activities relative to the overall budget of the Center, as follows:

- (a) # of epIA briefs published
- (b) # of disseminations of epIA findings in popular media
- (c) # of IA related conferences/workshops conducted for external audiences
- (d) # of IA related training materials developed
- (e) # of IA visiting specialists from a NARS hosted⁴

These figures were normalized in relation to a Center's overall budget such that those having one or more brief/event/etc. per \$10 million of Center budget (a SPIA-defined yardstick) received the highest possible score in each component and those less than that received a proportionately lower score.

The composite score for Criterion III was calculated by summing the scores from (a) – (e) above.

⁴ Another component, IIIA.3 (# of website hits) was eliminated due to the difficulty in interpreting the data across the Centers. Two other components (IIIA.4 – other dissemination; IIIB.4 – other activity) that had been listed were considered and evaluated by SPIA in conjunction with Criteria IV in computing an overall modifier score for each Center.

Overall Modifier: Additional points were given to Centers for other relevant information provided in IIIA.4 and IIIB.4 and in Criterion IV (Any other impact related activity). This was a subjective assessment made by all SPIA members.

Scores from Criteria I, II, III, and overall modifier points for each Center were tallied and summarized on an EXCEL spreadsheet.

Computation of the Final Score

The final Center scores for the 3a impact indicator were generated by summing up across all of the components of 3a (computed as described above) and the modifier points and rescaling the scores into a 1-10 scale by dividing by 10. The final scores are reported in Table 4.

Table 4. Center Scores for Indicator 3a - Commitment to Documenting Impacts and Building an Impact Assessment Culture

Center	Criteria*				Total Score	Total adjusted score (0-10 scale)
	I (max =70)	II (max=20)	III (max =8)	Modifier to overall score**		
Africa Rice	52.6	6.5	7.8	6.8	73.6	7.4
Biodiversity	21.9	15.5	6.3	3.8	47.4	4.7
CIAT	50.6	6.3	4.9	4.3	66.1	6.6
CIFOR	38.1	14.3	5.3	2.5	60.2	6.0
CIMMYT	47.5	10.9	4.5	7.3	70.1	7.0
CIP	50.1	14.8	3.6	1.3	69.8	7.0
ICARDA	48.5	16.9	7.1	6.3	78.7	7.9
ICRISAT	45.5	16.9	7.0	3.8	73.2	7.3
IFPRI	45.3	8.8	0.0	1.3	55.3	5.5
IITA	56.1	14.9	6.2	0.0	77.1	7.7
ILRI	34.2	10.7	2.5	1.3	48.6	4.9
IRRI	54.5	8.5	4.5	2.5	70.0	7.0
IWMI	21.8	11.7	2.8	1.3	37.5	3.8
World Agroforestry	41.2	13.1	4.2	0.5	59.0	5.9
WorldFish	52.7	10.4	6.7	0.5	70.3	7.0

* Criterion I = epIA studies/ advancement of methods; Criterion II = building an IA culture; Criterion III = communication/dissemination & capacity building; Criterion IV = other impact related activity

** The modifier to overall score is a combined score for criteria IV, IIIA.4 and IIIB.4.

3b. SC/SPIA Rating of Two Center Impact Studies Carried Out in the Period 2003-05 for Rigor

Information for this indicator is collected and evaluated by SC/SPIA once every three years. As this was done last year, the indicator rating carries over, i.e., remains the same as last year. There was brief discussion about the desirability and practicality of individual Centers

wishing to re-submit 3b IA case studies in the following year (e.g., for the 2007 PM exercise), especially in cases where they scored poorly. Given the magnitude of the task, SC/SPIA felt this was not feasible. Scores will therefore carry over until the 2009 PM exercise.

Annex I. Distribution of weights for 3a scoring exercise

	Component Weight	Indicator weights
CRITERION I. epIA studies/advancing methods	70	
I.B. EpIA studies		
1. Publication venue	10	
Refereed journal		10
Book chapter		7.5
Conference paper (includes proceedings)		5
In-house publication (reviewed externally)		5
In-house publication (not reviewed externally)		2.5
Unpublished report		1.25
2. (Co-) authorship	5	
With other CG Center scientists		1.67
With NARS scientists		1.67
With ARI scientists		1.67
Center only scientists		0.83
3. EpIA coverage	10	
Commodity improvement		2.5
NRM related		5
Policy related		5
Biodiversity related		5
Training/capacity building related		5
Other (specify)		5
4. Distance down the impact pathway covered by the study	15	
Uptake/adoption (field surveys)		3.75
Influence (bibliometric/citation analysis, key informant surveys, etc.)		3.75
Intermediate impacts (improved yield/quality, lower risk, higher income, conserve resources, increase market access/efficiency, develop human capacity)		7.5
Ultimate impact (poverty, food security, environment)		15
5. Geographical breadth of impacts assessed by the study	15	
Single location within single country assessment		1.9
Multi-locations (regions) within single country assessment		3.8
Multiple countries (~ 2-5) assessment		7.5
Global assessment (i.e., spread over several continents)		15.0
6. Advances in new methods/models for epIA embodied in the study	10	
Addresses non-economic impacts		2
Addresses differential effects (different target groups)		2
Addresses positive and negative effects		2
Addresses multiplier effects (other sectors)		2
Employs novel methods (combines quantitative & qualitative, participatory approaches, etc.)		1
Other methodological advances (specify :)		1

I.C. Annual Expenditure on epIA activity	5
Expenditure on epIA:	
Staff resources dedicated to epIA:	
CRITERION II. Building an IA Culture	20
II. A. List of Internal Workshops	5
II.B. Systematic Evaluation of User Relevance	5
II.C. Using epIA in Planning/Priority Setting	5
II.D. Baseline Studies	5
CRITERION III. Communication/dissemination & Capacity Build	8
III.A.1 epIA briefs	2
III.A.2 popular media	1.5
III.A.3 Website hits	not evaluated*
III.A.4 Other dissemination	modifier
III.B.1 Conferences/workshops	2
III.B.2 Training materials	1.5
III.B.3 IA visiting specialists	1
III.B.4 Other activity	modifier
CRITERION IV. Other Impact Related	modifier (+)
OVERALL TOTAL	98

* after lengthy consideration, SPIA decided not to evaluate this component this year given the ambiguity regarding its meaning. SPIA will be more specific regarding this component in next year's PM exercise.