

Defining and Refining Good Practice in *Ex-post* Impact Assessment – Synthesis Report

MARCH 2009



Defining and Refining Good Practice in *Ex-post* Impact Assessment – Synthesis Report

Meeting of CGIAR Center Impact Assessment Focal Points and the
Standing Panel on Impact Assessment of the CGIAR Science Council
held at the Alvaro Bracelos Conference Room, Embrapa Headquarters,
Brasilia, Brazil, 10-11 November 2008

MARCH 2009



The Science Council of the CGIAR encourages fair use of this material provided proper citation is made.

CGIAR Science Council (2009) *Defining and Refining Good Practice in Ex-post Impact Assessment – Synthesis Report* Science Council Secretariat: Rome, Italy.

CONTENTS

FOREWORD1

PROGRAM AGENDA FOR THE MEETING3

SUMMARY OF SESSION CONTENT AND OUTCOMES.....6

SESSION 1 - WELCOME6

SESSION 2 – INTRODUCTION BY DEREK BYERLEE6

SESSION 3 – CENTER PRESENTATIONS6

 Africa Rice – Presentation by Aliou Diagne6

 CIMMYT – Presentation by Roberto La Rovere.....7

 CIP – Presentation by Guy Hareau7

 IRRI – Presentation by Debbie Templeton.....7

 CIAT – Presentation by Doug Pachico7

 ICARDA – Presentation by Kamel Shideed8

 ICRISAT – Presentation by Jupiter Ndjeunga8

 IITA – Presentation by Victor Manyong8

 ILRI – Presentation by Nancy Johnson.....9

 Bioversity – Presentation by Elisabeth Gotor9

 CIFOR – Presentation by David Raitzer.....9

 IFPRI – Presentation by Tunis van Rheenen10

 IWMI – Presentation by Mark Giordano (via Tim Kelley)10

 World Agroforestry – Presentation by Steve Franzel.....11

 Harvest Plus CP – Presentation by Patrick Eozenou.....11

 PRGA – Presentation by John Dixon11

 Challenge Program for SSA – Presentation by Wale Adekunle12

 1) Policy oriented research impact assessment (PORIA)13

 2) Impact of agricultural research in South Asia since the Green Revolution (GR)13

 3) Strategic Guidance for Impact Assessment (SGIA).....13

SESSION 5 – DONOR DEMANDS FOR EPIA	14
1) ACIAR	14
2) Bill and Melinda Gates Foundation (BMGF).....	14
3) Embrapa	15
4) USAID/CIDA	15
5) Oxfam GB.....	15
SESSION 6 – STRATEGIC ISSUES RELATED TO EPIA IN THE CGIAR.....	16
1) Gaps in CG impact studies	16
2) CG Reform and its implications for epIA	16
3) Revised SPIA strategy	17
SESSION 7 - IMPACT INDICATORS 3A AND 3B.....	17
1) Modifying the Impact Indicator (3a and 3b): Proposal for discussion	17
2) Indicators for the magnitude of impact: Proposal for discussion	18
SESSION 8 – SOCIAL AND ENVIRONMENTAL INDICATORS	19
SESSION 9 – CLARIFYING PURPOSES AND EXPLORING NEW APPROACHES IN EPIA	20
1) Enhancing the Use of Ex-post Evaluation of Outcomes & Impacts in the CG	20
2) Institutional Learning and Change Initiative (ILAC)	20
3) Linking eaIA, epIA and priority setting	21
SESSION 10 - BREAKOUT GROUPS	22
1) Problem areas in epIA – <i>rapporteur Debbie Templeton</i>	22
2) Environmental and social impact indicators study – <i>rapporteur Paul Heisey</i>	23
3) Linking eaIA, epIA and priority setting – <i>rapporteur Jamie Watts</i>	23
SESSION 11 – OTHER TOPICS	24
1) Randomized Control Evaluation (RCE) for epIAs – <i>presented by James Stevenson</i>	24
2) Inter-Center collaboration, collectively moving forward on NRM IA and PORIA studies – <i>presented by Kamel Shideed</i>	24
3) Maintenance breeding.....	24
4) Under-explored IA areas – <i>presented by Greg Traxler</i>	25
5) Institutionalizing epIA beyond IAFP specialists – <i>presented by Tuenis van Rheenen</i>	25

SESSION 12 - CONCLUDING COMMENTS FROM THE SPIA CHAIR AND THE
PRESIDENT OF EMBRAPA26

ANNEXES27

Annex A: List of Participants.....27

Annex B: Proposed Indicator on Actual Impact29

FOREWORD

The meeting on '*Defining and Refining Good Practice in Ex-post Impact Assessment*' was jointly planned and organized by the impact assessment focal points (IAFPs) of the international centers and the Challenge Programs, and the Standing Panel on Impact Assessment (SPIA) as a follow up to a similar meeting held in Nairobi in 2006. The objectives of the meeting were to:

1. Foster communication among CGIAR centers on ex-post impact assessment
2. Contribute to a community of practitioners to share knowledge on impact assessment of international agricultural research
3. Receive feedback on SPIA ongoing activities
4. Make recommendations on future priorities for system level impact activities by SPIA or for joint impact assessments across centers

The meeting was successful in meeting these objectives due to the very active and constructive participation of 14 Centers, two Challenge Programs, two system-wide program initiatives, two donor agencies, several external evaluation experts, two SPIA members, three staff of the Science Council Secretariat and several scientists from EMBRAPA. In addition, one Center IAFP and two donors who were unable to attend sent along presentations for inclusion in the agenda. Special thanks are due to EMBRAPA for agreeing to host the meeting at the centre, and especially to Flavio Avila for their hospitality and taking care of all local arrangements. Tim Kelley provided invaluable support to the organizing team.

This Synthesis Report has been largely prepared by Tim Kelley, with input from James Stevenson and many participants. While it does not include any written papers, most PowerPoint presentations will be made available through the CGIAR Impact Website (http://impact.cgiar.org/documents/focal_point_group_meeting/) or from Tim Kelley (timothy.kelley@fao.org). This 'synthesis report' endeavors to capture the essence of the discussions and, where relevant, the consensus that emerged.

Derek Byerlee

Chairman of SPIA

PROGRAM AGENDA FOR THE MEETING

MONDAY NOVEMBER 10

- 8:15 – 8:25 **1) Welcome**
- Derek Byerlee, SPIA Chair
- 8:25 – 8:30 **2) Introduction**
- Meeting objectives and program agenda [D Byerlee]
- Logistics (F Avila/T Kelley)
- 8:30 – 9:45 **3) Center Presentations: Ex-Post Impact Assessment (epIA)**
- Each Center IAFP to make a 10-minute presentation followed by 5 minutes for questions; 15-minute general discussion at the end of each session
- A. Commodity Improvement Centers** [Chair: K Shideed]
- Africa Rice (A Diagne)
- CIMMYT (R LaRovere)
- CIP (G Hareau)
- IRRI (D Templeton)
- 9:45 – 10:00 Coffee Break
- 10:00 – 11:30 **B. EpIA at Commodity Improvement and Natural Resource Management (NRM) Centers** [Chair: G Traxler]
- CIAT (D Pachico)
- ICARDA (K Shideed)
- ICRISAT (J Ndjeunga)
- IITA (V Manyong)
- ILRI (N Johnson)
- 11:30 – 13:00 **C. EpIA at Policy, Institutions and NRM Centers** [Chair: P Heisey]
- Bioversity (E Gotor)
- CIFOR (D Raitzer)
- IFPRI (T van Rheenen)
- IWMI (M Giordano)
- WorldAgroforestry (S Franzel)

13:00– 14:00	Lunch Break
14:00 – 15:00	D. EpIA at Challenge Programs (CPs) and System-wide Eco-regional Programs (SWEPs) [Chair: E Gotor] Harvest Plus (P Eozenou) Sub-Saharan Africa (A Adekunle) PRGA (J Dixon)
15:00 – 15:15	Coffee Break
15:15-16:00	4) SPIA Update: Recently completed studies: [Chair: F Avila] a) Policy Oriented Research Impact Assessment – PORIA (T Walker) b) South Asia Impact Assessment – SAIA (T Kelley) c) Strategic Guidance for Impact Assessment – SGIA (M Maredia) Note: 15 minute discussion follows three 10 minute presentations
16:00 – 17:15	5) Donor demand for epIA – what are the expectations? [Chair: D Raitzer] Speakers: D. Templeton (ACIAR), G. Traxler (Gates), F. Avila (EMBRAPA), J. Stevenson (Oxfam/DfID), T Kelley (for USAID/CIDA) [40 min] Discussion: 35 min
17:15 – 18:30	6) Strategic Issues related to epIA in the CGIAR [Chair: D Pachico] a) Gaps in CGIAR epIA Studies (P K Joshi) b) CGIAR Reform and its Implication for epIA (R Echeverria) c) Revised SPIA Strategy (D Byerlee) Note: 15 minutes for each presentation followed by 30 minute discussion
18:30	Adjourn
18:45	Reception

TUESDAY NOVEMBER 11

8:00 – 9:00	7) Impact Indicators 3A and 3B in Performance Measurement System [Chair: D Templeton] a) Modifications to indicators 3A and 3B for the future (T Kelley) b) Indicators for the magnitude of impacts (M Maredia)
9:00-10:15	8) SPIA-led Social and Environmental Indicators Study [Chair: A Diagne] Speaker: Derek Byerlee - Discussant: Paul Heisey

10:15 – 10:30	Coffee Break
10:30 – 11:45	<p>9) Clarifying Purposes and Exploring New Approaches in epIA [Chair: T Walker]</p> <p>a) Enhancing the Value of Evaluations for Management study (Z. Ofir)</p> <p>b) Institutional Learning & Change (ILAC) initiative (Jamie Watts)</p> <p>c) Linking ex-ante (eaIA), epIA and priority setting (D Raitzer)</p> <p>Note: Each 15-minute presentation followed by 5-minute discussion with 15-minute discussion at the end</p>
11:45 – 13:00	10) Break out Groups for Items 7, 8 and 9
13:00 – 14:00	Lunch Break
14:00 – 15:00	<p>10) Break-out Groups Reports [Chair: T Kelley]</p> <p>Break-out Group #1: Social and Environmental Indicators</p> <p>Break-out Group #2: Impact Indicators 3A and 3B</p> <p>Break-out Group #3: Clarifying Purposes/New Approaches in epIA</p> <p>[10 minute presentation + 10 minute discussion for each]</p>
15:00 – 16:00	<p>11) Other Topics [Chair: D Byerlee]</p> <p>Randomized Control Trials for epIA (J Stevenson)</p> <p>Inter-Center collaboration/ Collectively moving forward on NRM IA and PORIA studies (K Shideed)</p>
16:00 – 16:15	Coffee Break
16:15 – 18:00	<p>11) Other Topics (cont.):</p> <p>CGI / maintenance breeding epIA (J Dixon)</p> <p>Unexplored IA areas, e.g. capacity building and germplasm collection and conservation (G Traxler)</p> <p>Institutionalizing epIA beyond IAFP specialists (T van Rheenen)</p>
18:00	<p>Closing Session</p> <p>Summing Up [D Byerlee]</p> <p>Final Remarks from Embrapa [Kepler Euclides Filho, acting Pres of Embrapa]</p>

SUMMARY OF SESSION CONTENT AND OUTCOMES

1 SESSION 1 - WELCOME

The SPIA Chair Derek Byerlee welcomed the participants to the 2008 IAFP-SPIA/SC meeting and thanked Embrapa for hosting the event, and the organizing committee including former SPIA Chair Jim Ryan for their hard work. In attendance were representatives or IAFPs from 14 Centers, two Challenge Programs, two system-wide program initiatives, two donor agencies, several external IA / evaluation experts, two SPIA members, staff of the Science Council Secretariat and several scientists from EMBRAPA (see Annex A Participants List).

2 SESSION 2 – INTRODUCTION BY DEREK BYERLEE

The SPIA Chair suggested a slightly revised program agenda from that previously distributed, which was accepted. He then reviewed the major objectives for the meeting:

- Foster communication among CGIAR centers on epIA
- Contribute to a community of practitioners to share knowledge on epIA of international agricultural research
- Receive feedback on SPIA ongoing activities
- Make recommendations on future priorities for SPIA system level activities or joint epIAs across centers

3 SESSION 3 – CENTER PRESENTATIONS

(Powerpoint presentations can be found at http://impact.cgiar.org/documents/focal_point_group_meeting/. This section merely highlights the topics addressed in each presentation)

Presentations were given by impact assessment focal points (IAFPs) from 14 of the 15 CGIAR centers and each of the four Challenge Programs. Each IAFP gave a 10-minute presentation as a quick overview of recently completed and on-going IA activities at each Center/Challenge Program, followed by an outline of future IA activities. Discussion sessions were held jointly, following presentations from three or four IAFPs at a time and chaired by a different IAFP. The issues raised and addressed in each discussion session can be grouped under four themes.

3.1 Africa Rice – Presentation by Aliou Diagne

Major topics included:

- Potential and actual adoption and impact of NERICA rice in West Africa
- Common methodologies across projects using household & plot level surveys
- Average Treatment Effect (ATE) and Local Average Treatment Effect (LATE)
- Impact of exposure and Impact of adoption in 5 countries in West Africa
- Estimating adoption “gap” of NERICA
- Impacts on yield, income, consumption spending, calorie intake, childrens’ schooling

3.2 CIMMYT – Presentation by Roberto La Rovere

Major topics included:

- Strategy to move towards multi-dimensional impacts, e.g., from crops to livelihoods.
- IA guidelines.
- Review of epIA, eaIA and priority setting studies, 1994 – 2007.
- Learning purpose of IA growing to complement the accountability function.
- IAs externally (1/3) vs. internally (CIMMYT staff) conducted (2/3).
- Future IA studies: Conservation agriculture; impact of impact assessment etc.
- Expansion of communication and capacity – progressive decentralization of IA work (Mexico, Ethiopia, possibly Asia).

3.3 CIP – Presentation by Guy Hareau

Major topics included:

- 2007 survey results: CIP derived varieties represent about 13% of total 8 million ha of potato area from surveyed countries, exceeding earlier estimates.
- Half of this is in China, other major countries include Rwanda, Peru, Uganda.
- Assumptions of adoption ceilings, now trying to understand factors explaining increased adoption levels.
- Potato's role in poverty alleviation; adoption leading to nutrition improvement.

3.4 IRRI – Presentation by Debbie Templeton

Major topics included:

- Numerous ('buried alive' in) impact-related initiatives.
- Trying to inculcate an impact culture.
- Evaluation is not just demand driven.
- Seven elements engendering real 'impact culture'.
- Instrumental role of MTP and change in PMI process.
- IRRI is more impact focused now because of: MTPs, PMI, strategic planning exercise, demand for impact in proposals, etc.

3.5 CIAT – Presentation by Doug Pachico

Major topics included:

- Technology assessment – accountability.
 - epIAs impressive (e.g., rice agronomy in Brazil and Venezuela), but only weakly influences donor decisions; and adds little knowledge
 - motivation is rather competitive – centre vs. centre
- Understanding complex innovations.
 - gender and social capital; environmental services studies;

- often yield results that are of more interest to donors
- hard to attribute/extrapolate, data and analysis intensive.
- Improving impact of research.
 - Pathway analysis & participatory monitoring methods
 - IA needs new methods to deal with complexity – output from PRGA
 - Huge potential for cross centre work on complex issues (role for SPIA)

3.6 ICARDA – Presentation by Kamel Shideed

Major topics included:

- Poverty impact of agricultural research – funded by USAID.
- Understanding perception and adoption of drought tolerant crop varieties.
- Cereals, food legume technologies in Nile Valley.
- Productivity and technical Efficiency.
- Mechanical water harvesting in Jordan.
- Investment in baselines, communication outputs.
- In-house seminars at centre-level.

3.7 ICRISAT – Presentation by Jupiter Ndjeunga

Major topics included:

- 72 programme areas vs. 71 counterfactual areas, using matching for sustainable land management project in Niger.
- IFPRI-ICRISAT-WorldAgroforestry – Adaptation to climate change.
- Multi-dimensional impacts of crop varieties – groundnut WAF, sorghum Nigeria, groundnut Uganda, meta-analysis.
- Secondary data at village level – matching done with propensity-score.
- ICRISAT contribution to SLM programme – credit attributable to ICRISAT.

3.8 IITA – Presentation by Victor Manyong

Major topics included:

- Impact studies on NRM and commodity improvement research.
- Working on building an impact culture.
- Cassava in Malawi.
 - Heckman’s treatment effects model for counterfactual analysis
 - Indicators: Yields, Per capita area planted, calorie intake in HH
- BNMS in Nigeria.
 - Two-stage least squares; survey on 400 HH exposed to technology
 - HH expenditure to food increased by 52% for BNMS-manure
- Impact of IITA in SSA on poverty and productivity (2009 in Food Policy).
 - Agricultural productivity growth, GDP per capita, etc. (secondary data)

- ROR 55%; CG research contributes 56% of the reduction in poverty

3.9 ILRI – Presentation by Nancy Johnson

Major topics included:

- Ex-post IAs.
 - Dairy policy impact in Kenya
 - IA of improved livestock-related soil management technologies in Ethiopia
 - Impact of improved animal health in the Philippines
- Re-thinking impact.
 - Experimental and quasi-experimental approaches (DfID cash transfer program in Kenya, conservation and use of endemic ruminant)
 - PRGA, ILRI and ILAC workshop in Cali this year
 - **How** research is done matters a lot, for **sustainable** poverty reduction
- Institutionalising impact.
 - IA task force for cross-theme group, working across disciplines/projects on standardising methods and data, linking knowledge to action
 - Experimental/quasi-experimental approaches for appraising research.

3.10 Bioversity – Presentation by Elisabeth Gotor

Major topics included:

- Previously, impact was just ‘assumed’ -- not necessary to analyse/document.
- Currently, designing programs from the beginning with IA in mind.
- Most work is through partnerships; so attribution of credit is problematic.
- CAB or CAE not always appropriate as the benefits Diversity produces cannot easily be monetised (as increases in yield or efficiency in input use can). Also, greater difficulty in identifying counterfactual scenarios.
- Requires new methods, e.g., for valuing agro-biodiversity.
- 10 areas of expected impact listed, e.g., “More supportive policy environments for the conservation and use of agro-biodiversity”.

3.11 CIFOR – Presentation by David Raitzer

Major topics included:

- One IRIS dealing with sepiia and priority setting.
- Recently completed studies.
 - Impact of CIPHER’s Criteria and Indicators Research
 - Uptake/bibliographic studies
 - Assessing impact of CIPHER’s influence on Indonesia paper & pulp sector
- On-going study: Impact Assessment of Network on Site Management and Productivity in Tropical Plantation Forests.
- Future Work: IA cases.
 - Contributions to forestry regulations in the Congo Basin (30m ha)

- Changes in community forest use rights in the Brazilian Amazon
- Forest use rights in Ethiopia
- Future Work: Factors affecting policy influence.
 - Scientist identify elements of research that led to policy influence
- Constraints and Issues.
- Looking to the future: ex ante assessment and priority setting.

3.12 IFPRI – Presentation by Tunis van Rheenen

Major topics included:

- Center IA focal point (Coordinator for Partnerships & IA) housed in DG’s office, but also focal points in each division.
- Case studies of epIA (all externally peer reviewed).
 - Ethiopia Strategy Support Programme
 - Less Favourable Areas (LFAs)
 - PROGRESA – one of the PORIA case studies
 - WTO’s Doha negotiations
- Looking forward.
 - Assessing impact at IFPRI through narratives more effectively
 - The impact of the conferences by 2020 Initiative
 - The Ghana Country Strategy Support Program evaluation
 - Global impact of IFPRI’s CCT research
 - Enhancing impact of IFPRI’s capacity strengthening program
 - Environment, Production and Technology Division evaluation
- Success stories from agriculture – “Millions fed”. Sponsored by Gates.

3.13 IWMI – Presentation by Mark Giordano (via Tim Kelley)

Major topics included:

- Impact planning a critical component of each theme in new strategic plan.
- IA elevated to a formal program in the institute’s research structure.
- Senior IA economist being hired mid 2009 for:
 - epIA and eaIA of IWMI and non-IWMI water related interventions
 - non-economic analyses of selected IWMI project and program impacts
- “Impact team” involved in all new project proposals / workplans
- On going epIAs:
 - Groundwater governance through electricity supply management:
 - Multiple Use Water Services Project
 - Economic Benefits from Improving Soil Fertility and Water Holding Capacity with Clay Application
 - IWMI gender and water programs

World Agroforestry – Presentation by Steve Franzel

Major topics included:

- Major impact studies (on-going)
 - Improved shrub fallows (Zambia and Kenya)
 - Fodder shrubs (East Africa) - gender analysis
 - Women plant fewer shrubs – risk aversion?
- A promising innovation (future impact?): monitoring changes in soil quality using light reflectance spectroscopy.
- Key challenge: how many farmers are using a new agroforestry practice?
- Pseudo-adoption: the adoption of a practice not for its intrinsic value in production but to accrue benefits from organizations promoting the practice.

3.14 Harvest Plus CP – Presentation by Patrick Eozenou

Major topics included:

- Existing impact studies (ex-ante).
- Key impact indicator: Disability adjusted life year (DALY); cost per DALY is below \$200.
- Methods being used:
 - consumer acceptance survey
 - willingness to pay survey
 - simulations (impact) – partial equilibrium model
- Forthcoming impact studies (ex-post).
 - Orange Sweet Potato in Uganda and Mozambique
 - prospective randomised control experiments: control, model 1 (intensive), model 2 (less intensive)
- Adoption rates, yields, profitability and commercialisation.
- Treatment package: agricultural extension component (distribution of vines and training and demand creation component (nutrition messages with special emphasis on health related benefits).

3.15 PRGA – Presentation by John Dixon

Major topics included:

- New strategy – reorientation and shift of emphasis.
- PPB and seed delivery systems.
- Measuring effectiveness of research processes in complex systems.
- Gender mainstreaming.
- Complexity workshop in Cali.
- Poverty - Multi-dimensional and complex.
- RoR studies not well suited to complex systems.

3.16 Challenge Program for SSA – Presentation by Wale Adekunle

Major topics included:

- Research on interfaces of productivity, NRM, markets and policy.
- Approx half of the 80 partners involved are from civil society.
- Outcomes – Increased capacity, increased diffusion and adoption, understanding of impact pathway under institutional innovations etc.
- Key Question: Does IAR4D work and does it deliver IPGs?
- IAR4D villages.
- Before and after – With and without.
- Counterfactual villages 1 – No external input from research.
- Counterfactual villages 2 – External inputs pre-existing IAR4D.

Discussion points from Center presentations:

- As it was referred to a number of times in the presentations, some questions arose concerning the definition and meaning of the “learning function” of epIA. Many agreed that there is a need to emphasise learning drawing on not just a single study but on meta-analysis. There are risks in drawing conclusions from a single study which are subject to time and spatial contextual constraints.
- As Centers like CIMMYT and IRRI have had more (documented) impact than other centers over the past 15 years, they could perhaps lead the way in systematic efforts to document impact further down the pathway in the future IA.
- Need to make a distinction between centers’ activities targeting analysis of the impacts of non-research interventions, including of natural phenomenon, e.g., drought, vs. documenting impacts of research per se.
- Impact of IA on donor behaviour: seems not to have played a role in resource allocation; however, at centre level (CIP), IA has played a role in divesting from areas that did not demonstrate impact.
- Best predictor of the future is understanding history. Ideally, want to draw conclusions from epIA as a basis for resource allocation, i.e., importance of empirical evidence. But one also needs to understand the processes/ historical analysis; not that you can’t extrapolate, but you need detailed historical analysis to inform assumptions for what you can achieve in the future.
- (Value of) meta-analysis depends critically on having a body of individual centre studies (re: SPIA’s role in meta-analysis studies). The importance of having coordinated efforts to identify studies around particular important studies – gender / empowerment for example, was emphasized.
- Priority of IA of non-research activities? Working out where we are on the continuum from research to advocacy. Comparative advantage issue: finding a partner organisation who has the skills and reach to do some of this.
- What type of studies should Centers be doing internally vs. commissioning externally? Depends on the nature of the study, skills and time available in the

- Value of understanding contextual factors that determine success may sometimes be a better research question than seeking to determine what has been the impact.

4 SESSION 4 – SPIA UPDATE

4.1 1) Policy oriented research impact assessment (PORIA)

Tom Walker provided a brief summary of the motivation, objectives and key results of Phase 2 of the PORIA study which involved seven case studies designed to document the influence and measure the impact of POR. The final report and brief and seven case study briefs have just been printed. Some selected highlights included:

- Confirms Timmer’s hypothesis, i.e., the need for long-term presence / association to have impact.
- Discounted net benefits for five of the case studies plus three earlier studies that had estimated impact comes to just under \$1 billion, just about breaking even on total CGIAR investment in POR.

Issues discussed included: appropriate methods for establishing Center influence (key informant surveys were mainly used); major purpose of the study – largely accountability: donors asking for evidence of POR impact at AGM four years ago.

4.2 2) Impact of agricultural research in South Asia since the Green Revolution (GR)

Tim Kelley provided a brief summary of the Peter Hazell report, commissioned by SPIA to address a number of concerns about the purported negative impacts from the agric research in post GR South Asia. Thus, this comprehensive assessment was undertaken to better understand the direct and indirect pathways of past impacts (economic, social and environmental, positive and negative) of CGIAR + partner research on different producer and consumer groups, and to draw out key lessons.

The main findings of the study: (i) the CGIAR + national partners have responded well to post GR criticisms and to agriculture’s evolving needs; having shifted from a narrow focus on productivity improvement to goals of poverty reduction and enhancing the sustainability of natural resources, the CG’s work continues to be a sound investment for the region; (ii) agricultural research has reduced poverty through growth and food prices, but has been less effective at reducing inequalities; (iii) science has also contributed to reducing environmental damage via watershed development, more efficient use of fertilizers and water, etc. But the report highlights the scarcity of empirical studies that trace causal connections between investments in research and measurable outcomes in terms of poverty. In the discussion, a question was raised about how much influence this and the PORIA study reports have on donors, i.e., what has been the impact of these impact assessments? What is the strategy for communicating and for what effect, and purpose? In some sense, it’s comparable to maintenance research -- if you don’t do it, you will see the (negative) effect, and donors will go elsewhere to get the information. SPIA agreed more thought needs to be given to communicating results effectively.

4.3 3) Strategic Guidance for Impact Assessment (SGIA)

Mywish Maredia presented an update on this SPIA-led initiative, which has just been completed and a report (and brief) published. She described the rationale and objectives and

some of the key elements of this strategic guidance document that was developed in collaboration with Center IA focal points. The document proposes a clear and definitive definition of epIA and their purpose. EpIAs should be viewed as a two-stage process consisting of measurement of either or both (i) direct effects on beneficiaries (usually adopters/producers) and (ii) indirect, often multi-dimensional impacts on a range of stakeholders (producers, consumers, laborers). The discussion that followed focused on several issues, including:

- The share of resources that should be allocated to assessing impacts – up to 3% is suggested in the SGIA, not all of that is epIA (eaIA).
- The extent to which we can monitor web use and uptake of these guidelines and how to inform/disseminate to NARS.
- How to keep this guidance document ‘evergreen’ – a suggestion was made for setting up a Wikipedia like site for this.
- The degree to which the guidelines are realistic and helpful given some major constraints for documenting impact (highlighted in Bioversity’s presentation).
- How this guidance document lines up with other evaluation agencies standards, e.g. NONIE.

5 SESSION 5 – DONOR DEMANDS FOR EPIA

5.1 1) ACIAR

Debbie Templeton presented ACIAR’s perspective on impact assessment of its research investments. ACIAR’s IA program currently undertakes two types of finished project assessments: adoption studies and *ex-post* impact assessments.

As of today, over 50 full benefit–cost assessments had been published in ACIAR’s Impact Assessment Series, and 49 adoption studies had also been undertaken.

Adoption studies are undertaken on completed projects where ACIAR expenditure is greater than A\$ 400,000, 3 years after the completion of the project (where there is no follow-on project). Motivation for these studies: (i) Accountability; (ii) Lessons learnt – which feed into in-house reviews; and, (iii) Methodological advances.

5.2 2) Bill and Melinda Gates Foundation (BMGF)

Greg Traxler outlined the major considerations BMGF uses in evaluating proposals for funding. They put heavy emphasis on establishing a clear linkage between research/development activities and improvements in key indicators of human well being. The latter relate to: household income, household food consumption and child nutrition status. In recognition of that long run impacts are difficult and costly to measure, and attribution not easy to establish, BMGF relies on intermediate indicators (e.g., value of HH assets, income from agriculture, value of farm product sales in market) and puts more emphasis on baseline surveys, project monitoring and ‘theory of change’. The M&E strategy for the AgriculturalDev Unit (within Global Dev Division) is still under development. The 3 divisions (Global Health, Domestic Programs, Global Development) compete for funds – the co-chairs look for results against core objectives. The planning stage requires a vision of success, i.e., a description of the desired state of the field in the specified number of years of the project. At the end, must tell a credible story of how BMGF funds have contributed to a

transformation of agriculture and impacted on human welfare. The demand is for IAs that can show improved quality of life. The minimum requirement is to show plausible influence on one of the 3 core indicators. Internal rates of return (IRR) is, at present, not used in BMGF. The CGIAR system is 2nd largest grantee in Agricultural Development.

5.3 3) Embrapa

Flavio Avila presented EMBRAPA's perspective on demand for impact assessment. Demand at the end of the 90s was for multi-dimensional evaluations. New System developed called Social Balance.

5.4 4) USAID/CIDA

Tim Kelley highlighted a few key points from a paper prepared by Meredith Soule (USAID), Saharah Chapotin (USAID) and Iain MacGillvray (CIDA).

- Research is risky and we expect to have numerous research programs that do not yield significant impact on the poor; but we expect some successes, and past experience has shown that major successes yield impacts that are large enough to pay for those and the 'non-successful ones'.
- For programs that do not have measurable impact, must distinguish between poor implementation and design and research that did not deliver anticipated results; lessons from the former should be used to improve the institutional process and lessons from the latter will inform future research and program decisions.
- For the CGIAR there have been two key annual management decisions for donors: 1) how many \$ to allocate to CGIAR system from scarce development resources; and 2) how to divide those resources among the 15 CGIAR centers. For the first, results achieved in the past (via ePIA studies on rice, wheat, cassava) have been valuable in keeping a focus on agricultural research within the overall development portfolio. Another big breakthrough would give a boost to the arguments for the continued relevance of international agricultural research.

5) Oxfam GB

James Stevenson outlined Oxfam's experience of reporting to donors, with a focus on the relationship with DFID. DFID's resourcing to International Non-Governmental Organisations (INGOs) has come under scrutiny in recent years from the bodies that monitor value for money in the use of tax-payers money in the UK. The agreements to fund development programmes through INGOs have been poorly evaluated for a long time, with little academically credible evidence of impact on poverty alleviation presented across a whole cohort of organisations. Oxfam and DFID have a shared interest in putting this situation right, and are now working on mechanisms that meet accountability concerns on both sides. Oxfam's current approach focuses on repairing three critical areas of weakness:

- Clearly articulating *ex-ante* impact pathways.
- Defining appropriate indicators for poverty outcomes and impact.
- Resourcing rigorous data collection.

DFID have request a series of baseline studies in Oxfam programmes, to monitor change over the next three years. In 2010/11, the collective portfolio of evidence across a group of

UK-based NGOs will be used to determine future funding levels to these organisations. While these system-wide initiatives take time to become business as usual, Oxfam will provide some financial resources (and look for new sources of research funding) for a limited number of more detailed impact evaluations for specific types of programming.

General discussion (consensus) points:

- Demand for IA is rising, but at what level is this having an ‘impact’? Widely held view is that if the system wasn’t producing enough credible impact studies, the budget could not be sustained”. They play an instrumental role according to USAID/CIDA. ACIAR’s funding has risen primarily due to the belief that Centers are a good investment; IAs have been used to justify the increase. Methodological advances from the CGIAR have also been important to them.
- On the question of whether donors consider size and scale of impact, all BMGF projects are using impact indicators and in theory are looking for large impacts, but it’s still early days (i.e., not very sophisticated yet).
- Does the CGIAR itself uses IA (a donor asked)? If so, why has crop improvement activity been marginalised, despite the substantially demonstrated high returns, whereas resources for NRM and POR have increased with little evidence of impact? Surely Centers have some scope for allocating unrestricted resources.
- Reference made to a key finding from the Donor Demands for IA study in 2005: linking evidence from past impact to future resource allocations – no apparent relationship, i.e., other factors are guiding choices.

6 SESSION 6 – STRATEGIC ISSUES RELATED TO EPIA IN THE CGIAR

6.1 1) Gaps in CG impact studies

P K Joshi who participated as a consultant in the Independent External Review of the CGIAR gave a presentation highlighting areas of research in the CG that are under-assessed with respect to impact, and require attention, especially NRM and gender related research. There are also a very limited number of studies that document environmental and social impacts from CG research, and few macro-level impact studies, noting the shift from a global to local focus for many Centers. The CG needs to give more attention to scaling up impacts on socio-economic and environmental indicators.

6.2 2) CG Reform and its implications for epIA

Ruben Echeverría elaborated on recent developments with respect to the current CGIAR reform process and implications for IA in the system. These reforms have the potential to strengthen epIA. Although exactly how IA will be institutionalized is not yet known, having the Consortium should result in greater coherence and coordination of IA internally, but the need for external IAs remains. Performance contracts may stipulate periodic epIA for successful research outcomes. There is likely to be more emphasis on social and environmental indicators of impact, and particularly examining impacts on poverty reduction. Continuing to improve in a cumulative manner (system-level) what SPIA has always done was highlighted, and these require global, regional and sub-regional type analyses. There will be more attention to agricultural innovation systems and measuring impact of those. The IAFP community of practice should remain and new collaborative arrangements with groups outside the CGIAR with expertise in IA (NONIE, universities)

forged. The three options for epIA in the future: (i) centralized panel of experts; (ii) strengthen SPIA (iii) outsourcing.

6.3 3) Revised SPIA strategy

Derek Byerlee presented the main elements of the current (recently revised) SPIA strategy, but highlighted that the strategy may need significant revision in light of the outcome of the CGIAR reform process. Results from this meeting itself are also likely to feed into a revised strategy. The three objectives of SPIA were reviewed (accountability at System level, strengthening Center IA, facilitating feedback). SPIA/IAEG has been guided by: (i) focus on impact (not adoption); (ii) *ex-post*; (iii) accountability to donors; (iv) operating at System level; (v) adding value to what Centers do (90% of epIA activity). Some issues requiring reflection were put forward for discussion: What has been the impact of SPIA and how do we assess its effectiveness? How useful (or onerous?) is the CGIAR impact website; who are the primary users—donors, NARS or IAFPs? Should SPIA continue to develop/maintain it? How important is individual attribution of impact in this age of increasing partnerships? Should SPIA and IAFPs interact more with other evaluation groups?

General discussion points:

- Need to better link epIA results with eaIA (with growing importance of the latter) but mixed views on whether SPIA should take on eaIA at system level.
- It's not just about filling gaps but the need to assess priorities for epIAs, e.g., update earlier CGI work where arguably biggest successes lie, what's the feasibility of addressing gaps (NRM, training), where are the payoffs highest?
- With wider and stronger partnerships in the new CGIAR, likely that we will move beyond attribution of quantitative of credit to individual centers – “joint products”.
- Agreed with need to work more with other evaluation groups (NONIE, 3IE) with opportunities to contribute to their methods and vice versa.
- With the mega-programs, more emphasis will be on system wide impacts, but it was not clear how the entire M & E function will operate in the new model, i.e., internal vs. external instruments.
- Growing importance of results-based programming – and developing impact indicators for each strategic objective and setting targets for overall CGIAR system in line with MDGs (guiding ODA community and countries).
- A strategy is needed for strengthening NARS capacity with respect to IA.
- IAFPs are likely to be strongly affected (i.e., burdened) by the changes implicit (transactions costs) in the reform process.

7 SESSION 7 - IMPACT INDICATORS 3A AND 3B

7.1 1) Modifying the Impact Indicator (3a and 3b): Proposal for discussion

Tim Kelley introduced the first part of this session by presenting a proposal for discussion that was developed by SPIA based on recommendations from a workshop on PMS in Washington D.C. in July 2008. The major changes proposed (from last year) included the following:

1. Merging 3A and 3B components into single indicator “Impact Culture”.
 - i) epIA studies and new methods (45%)
 - ii) IA culture and capacity building (20%)
 - iii) epIA rigor and quality (35%)
2. Terms clearly defined (e.g., what is an epIA?) and rationale for impact culture indicator explicit.
3. Consolidating various internal components (simplified).
4. Re-assessing and modifying weights of components, including more prominence (higher weights) to ‘negative effects’.
5. Lowering the benchmark for optimal number of epIA submitted, from 1 epIA per \$5 m to 1 epIA per \$10 m of Center budget.

There was a consensus about the need to lower the benchmark, as the current one creates a moral hazard – encouraging Centers to generate many small scale adoption type studies rather than documenting high quality epIAs that demonstrate large scale direct and indirect impacts on a range of economic, social, environmental indicators. In a follow-up meeting with the IAFPs, agreement was reached on lowering the benchmark even further (for reasons cited above) to 1 epIA per \$20m.

7.2 2) Indicators for the magnitude of impact: Proposal for discussion

Mywish Maredia introduced the second part of this session – elaborating on 3 alternative proposals for developing and applying an indicator of ‘actual impact’ of research done by a Center, i.e., going beyond impact culture:

1. Periodic meta-analysis at the center level;
2. Periodic impact assessment of Centers based on in-depth analysis of “impact claims” – the case study approach; and,
3. System-level meta-analysis on an annual basis.

The three proposals (described in Annex B) generated much discussion, and concern. Generally, it was felt by both SPIA and the IAFPs that virtually all three proposals would be excessively data and time intensive and would require careful planning and resources for execution. A major constraint related to the rules of aggregation. Moreover, given the heterogeneity in the type of research done by different Centers, it was felt that it would not be feasible to come up with a single method or consistent set of indicators of actual impact applicable to all Centers. A point was made about what is actually being measured and evaluated by donors in this exercise: most epIAs are measuring impact from research done 15-20 or more years ago, so this is not a useful exercise in terms of measuring (and comparing) current impact. There was also agreement that with the anticipated structural changes about to take place, the timing was not right for introducing a new indicator into the PMS. Taking all of this into account, and particularly noting the technical challenges (no other research systems attempt this), the group concluded that the issue of an indicator of “actual impact” should be re-visited only after the new CGIAR system is organized and functional. It also felt that a more viable and realistic option to explore in the longer run is to encourage Centers to compile and synthesize documented evidence (both quantitative and qualitative) of actual impacts every five years prior to and as an input into their scheduled

external program and management reviews.

8 SESSION 8 – SOCIAL AND ENVIRONMENTAL INDICATORS

Derek Byerlee presented an update of a SPIA commissioned study to explore greater use of environmental and social indicators (ESI) of impact in the CGIAR. Donors want to move beyond directly measurable economic impacts to account for ESI impacts—negative and positive. Joint SPIA/IAFP work has evolved to consider NRM research and poverty impacts but is still heavily focused on economic impacts expressed through markets, with a few important exceptions—pesticides and health, carbon emissions. This activity emerged from the IAFP meeting in Nairobi which confirmed the need for exploring and applying ESI using new approaches. This study will be employing two broad approaches:

A. Extended Cost-Benefit Analysis (ECBA) expanded to include non-market ESI where all benefits/costs are converted into monetary terms, using revealed preferences (hedonic pricing) and stated preference (contingent valuation, choice modeling). Jeff Bennett at ANU is the consultant.

B. Indicators and other measures of ESI that are considered separately with market costs and benefits using weighted indices, cost effectiveness, etc. (consultant: tbd).

There are three stages to the study: (i) Desk review of methods and identify relevant case studies; (ii) Case studies led by centers; (iii) Synthesis of lessons and development of guidelines. We are currently in the early part of Stage i.

Paul Heisey offered the following reflections on the proposed study:

- Another method for observation-based measures of environmental costs (or benefits) is changes in production costs, i.e., observing changes in firm profits, input costs, or output prices due to changes in environmental quality.
- Considerable care and expertise must be taken to apply methods for measuring the economic values of non-market goods or bads, even when these methods are applied in comparatively simple valuation exercises. When measurement is undertaken in conjunction with analysis for which attribution is a major issue (e.g. research impact or policy impact), the chances of compounding mis-measurement are correspondingly greater.
- In general, measurement of non-market goods or bads has not been formally integrated with research evaluation.

Key observations arising from the discussion included:

- Most 'environmental research' (especially CIFOR, WorldFish, IWMI) is actually policy research. The key issue is not so much valuation, but attribution.
- Risk of double-counting benefits, e.g., embedding soil erosion benefits in the biomass production.
- The importance of measuring changes in key variables in their own units, even if one is not able to value them, i.e., distinguishing between problems associated with quantification and valuation. Agreed that quantification is the most critical step and there was a dearth of studies that have done even that.
- This study should help us in providing guidance for valuing genetic resources, its

conservation value and use value.

- Agreement that we need to go beyond CBA in this study.

9 SESSION 9 – CLARIFYING PURPOSES AND EXPLORING NEW APPROACHES IN EPIA

9.1 1) Enhancing the Use of Ex-post Evaluation of Outcomes & Impacts in the CG

On behalf of Zenda Ofir (SPIA focal point on this study), *Sirkka Immonen* presented a summary of the recently completed desk study on the above topic. The demand for this study derives from an increasing emphasis given to learning for decision making and priority setting from all forms of evaluation, of which epIA is one. SPIA maintains that learning should not come at the expense of de-emphasizing accountability. This study therefore was intended to look at (a) the manner and extent to which the evaluation results, particularly epIA, can be utilized for learning, and (b) how the *approach* taken to such evaluations, including process, enhances or does not enhance the learning that takes place. A review of the broader literature documenting the role of ex-post evaluation for learning, decision-making, and other use and the 8 interviews of selected informants were used to identify the four major kinds of evaluation use: accountability, conceptual, instrumental and process. The different kinds of evaluation information from M&E, outcome evaluation and epIA are more or less useful at different organizational levels be it project, program, centre, system or donor. With a growing tendency to advocate for certain types of methods, there is a need to rigorously test promising practices for appropriateness.

9.2 2) Institutional Learning and Change Initiative (ILAC)

Jamie Watts described the objectives and major activities of the ILAC initiative. ILAC is currently focused on lesson learning and experience sharing among Learning Lab members, field testing approaches and methods for collaborative agricultural R&D, developing or testing approaches and methods for evaluation and providing training workshops and technical assistance. A major project underway, “impact evaluation” seeks to improve impact evaluation of collaborative, applied agricultural research programs aiming for poverty alleviation and add to the toolbox of impact evaluation methodologies to address known gaps. A major component of this project is carrying out a set of exemplary evaluations that develop methods and synthesize results recognizing the complexity of the impact pathway. Within this context a number of key preliminary questions will be addressed, e.g., how can the evaluation process engage expertise, methodologies and perspectives across a variety of relevant disciplines? The ILAC initiative questions accountability imperative – to whom and for what?

Patricia Rogers followed up with a brief summary about using epIA for learning:

- About: what works, what works for whom in what contexts, what doesn't work.
- By whom: project, program, institution, partners, donors, etc.
- From: findings, process – articulating impact pathway, developing baselines, collecting contextual and process data, negotiating with partners, etc.
- Of: individual studies, synthesis of multiple studies – meta-analysis.
- Through: carefully designed process, serendipity.
- For: priority setting, project design & management, future eaIA and epIA.

And, using epIA for accountability:

- Who is required to account: Programs, centres, CG system?
- To whom: Doners, NARS, partner organisations?
- For what: achieving intended impacts, building new knowledge, etc.
- Through what means: External IAs, longer-term follow-up, etc.
- With what consequences: Changes in investment, personnel changes?

9.3 3) Linking eaIA, epIA and priority setting

David Raitzer argued in his talk “The missing link: ex-ante IA to bridge ex-post findings with future priorities” that epIAs are a blunt instrument for learning. Although there is now increased demand for ‘learning’ from evaluations, long lag times mean that epIAs offers insights that are only indirectly related to present research choices (due to adoption context, social needs and research capacity all evolving).

Some key points made:

- Donor perceptions of future impact diverge from epIA results.
- Limited role for considering past impacts in selection of research priorities.
- Need for *ex ante* impact assessment and priority setting.
 - to make tacit assumptions explicit
 - to test assumptions against history
 - higher probability of having impact ex-post if ex ante is conducted – Jeff Davis, ACIAR
- Lack of coordination or guidance so far in the CGIAR, but some strong examples of eaIA activity at center level.
- CIP priority setting (eaIA) exercise offered a rigorous approach to learning, including divesting of work that did not look promising.
- Learning process only possible when explicit assumptions are outlined.
- Increased attention to defining impact pathways – Identification of the 5As:
 - Aim – what is the policy that needs to be changed
 - Ammunition – evidence for change that is required
 - Actors – Those who influence
 - Allies – Those likely to use the ammunition
 - Plan of Attack – How the research can be used by the allies

Discussion points from the presentations

The general discussion that followed touched on a wide range of issues. The Chair of the session noted a “dissonance problem” as evident from results from the Donor Demands for Impact study (2005), with documented impacts not aligning with perceived impacts and expectations of future impact not aligning with past impacts. SPIA has a role in correcting such mis-perceptions.

Other issues discussed included:

- The need for making critical distinctions between scope for learning from different types of evaluation across the spectrum, from ex-ante IA (high) to epIA (low).
- Insufficiently addressed are site specificity factors and contextualization (both in space and time) – both are critical for putting boundaries on the ‘learning’.
- “Intelligent failures” as used in the presentations, conveys the wrong idea; when results from research are used to reject a null hypothesis, for example, this is not failure-- in terms of building a body of evidence concerning certain hypotheses.
- Need for conceptualizing the different evaluation processes used in the CGIAR.
- Selecting candidates for epIA, e.g., innovation platform in SSA CP.
- Ex-ante IA have a key role to play in enhancing credibility of proposals (underpinning key assumptions about the potential magnitude of success expected and the target group).
- GIS specialist work within eaIA modeling has been (and remains) valuable although needs more use of hard evidence/data rather than ‘expert opinion’.
- Still too much focus on B-C analysis and too little on users, key assumptions.
- Key question raised by a donor: in the focus on ex-ante scenarios in the (IFPRI) modeling effort, how much real monitoring is done... do we know the main indicator of diffusion / adoption at the outcome level?

10 SESSION 10 - BREAKOUT GROUPS

Three break-out groups were formed to delve into more depth on issues of interest to the wider group. The three issues of most relevance to the group were: (i) problem areas in conducting epIA; (ii) environmental and social indicators of impact; and (iii) linking eaIA, epIA and learning. Groups were self selected and after a two-hour break-out the rapporteur from each group gave a short synopsis of the discussion.

10.1 1) Problem areas in epIA – *rapporteur Debbie Templeton*

Within the context of the Strategic Guidance session discussion, i.e., issue of whether the guidelines are really helpful or realistic, five major challenges in conducting epIA were highlighted during the break-out discussion:

- Counterfactual / attribution – major stumbling block especially for difficult areas like policy and advisory (Bioversity’s work).
- Lack of baseline data / lack of rigor in data collection (surveys expensive).
- Lack of skills (internal or external) to carry out the epIA rigorously.
- How do epIA results truly inform investment decisions – how can it be made more utilization focused?
- Demands for multiple types of information from multiple stakeholders.

A question was raised about the need for attempting attribution – is it really necessary to assign credit among your partners? It may not always be, but often donors want to know

what the return on their investment has been. The most important part is establishing the counterfactual, i.e., what would have been the situation without the project. In some cases, e.g., when the counterfactual represents the case without the CGIAR center's input or influence, attribution is derived directly from comparing the case of the current situation with the counterfactual.

10.2 2) Environmental and social impact indicators study – *rapporteur Paul Heisey*

- The discussions focused on several key points:
- Clear recognition that more sophisticated methods for monetizing indicators is needed, but must be cautious about this. There is also a need for methodological innovations in getting raw quantitative data, even in physical terms.
- There is an impressive amount of experience within the system in measuring social and environmental variables – need to draw on those individuals. But there are methodological issues related to extrapolation from HH to village to regional levels.
- Need for building on and tapping into relevant ARIs who have done considerable work in this area, methodologically and empirically.
- Question of whether there should be one or two groups (one for social indicators, one for environmental) – they interact but are separate issues.

In the plenary discussion it was suggested that SPIA should share the scoping study report with the IAFPs to get their feedback prior to finalization. It was pointed out that there are a few examples of measuring impacts related to the environment (CIFOR policy work on deforestation). A suggestion was made that for the Phase II case studies, it would be useful to build on earlier economic B-C IA studies – in trying to build out more comprehensive IA analyses that incorporate social and environmental impacts.

10.3 3) Linking eaIA, epIA and priority setting – *rapporteur Jamie Watts*

The group identified a number of principles relevant for enhancing this linkage:

- Recognize and manage learning and priority setting at different levels, e.g., at project, program, MTP, CGIAR levels.
- Priority setting should be more bottom up than top down in planning process, drawing extensively on scientists, field staff, etc.
- For the SPIA commissioned enhancing learning study, more manageable TORs are required if it's to be successful.
- Need an inventory of where the largest knowledge gaps are.
- Knowledge areas for investigation could be processes (e.g. collaboration) or thematic.
- A major weak link is the absence of epIA results in eaIA priority setting work.

An issue raised during the plenary was whether there exists the capacity to change once lessons are learned. With reference to the SPIA learning study, the feasibility of moving to a second phase was questioned given the timing (major reform process on-going), but if it did it required more structure, i.e., there is a need to construct and test hypotheses about epIAs rather than just assembling information from various individuals – who vary considerably in their experience of conducting/using epIAs.

11 SESSION 11 – OTHER TOPICS

11.1 1) Randomized Control Evaluation (RCE) for epIAs – *presented by James Stevenson*

RCEs use experimental methods to eliminate ‘selection bias’ between treatment and control group through random assignment. They are being increasingly used for their strong counterfactual and high internal validity, i.e., they depend less on assumptions regarding steps along the impact pathway. The method appeals to donors (Oxfam is supporting some RCE efforts in projects in Kenya) and many academics in development economics, but RCEs are controversial and criticized by many in the evaluation field. Whether they are useful for documenting impact (as opposed to simply establishing efficacy) may depend on validity of assumptions about or ease of scaling up, type of intervention considered -- simple vs. complex. Other concerns relate to cost – RCEs are expensive to implement (though cost of a wrong answer is very high). Arguments in favoring of exploring or expanding use of RCEs in the CGIAR include: i) Centers are already doing them; ii) scope for designing new experiments to address economic, social and environmental impacts together exists (though expensive); and iii) there are new sources of funding for this approach and new venues for reporting “intelligent failures” to overcome publication bias. A major comment during plenary discussion focused on applicability – the approach seems more relevant for establishing efficacy in a limited environment than one for documenting impact *ex-post*. SPIA will be commissioning a scoping paper to explore this issue further.

11.2 2) Inter-Center collaboration, collectively moving forward on NRM IA and PORIA studies – *presented by Kamel Shideed*

Due to the inherent complexities in documenting impacts from NRM and policy research, multi-centre approaches and interaction are deemed highly desirable. Three major rationales were given: (i) economies of scale – levels of expertise required are often beyond capacity of individual centers, requires critical mass; (ii) synergies and complementarities, based on comparative advantage; and (iii) enhance exchange/flow of information, methodologies and experience. Four priority areas of inter-center collaboration were highlighted (of 10 listed): (i) social and environmental IA, especially quantitative and qualitative indicators and valuation techniques; (ii) macro-level impact studies; (iii) methodologies for NRM and policy research (trade-offs, scale, establishing ‘influence’); and (iv) communication and utilization of IA studies (spillovers). Two research theme examples (PPB and conservation agriculture) were given as ways in which added value could be given by using a multi-centre approach to documenting impact. In the discussion that followed a few other topics were suggested as candidates for IA collaboration – ex ante IA, ILAC, CAC systemwide program, and one centre (CIMMYT) has already established a five-year plan for IA activities – this could be a basis for inviting collaboration. IA collaboration across Centers doesn’t necessarily need SPIA involvement.

11.3 3) Maintenance breeding

John Dixon spoke briefly to the issue of IA of maintenance research. An ex-ante assessment has already estimated a value of \$1 to 8 billion dollars for maintenance breeding for wheat leaf rust (to control UG 99). But there are very few epIAs of maintenance breeding, yet it constitutes a significant portion of research investment. He suggested this would be a good candidate for a SPIA commissioned IA meta-analysis.

11.4 4) Under-explored IA areas – presented by Greg Traxler

There appears to be sparse evidence of impact of NRM/crop management research compared to crop improvement. Is it due to (a) lack of impact (many studies showing low returns?) or (b) lack of effort to document impact (few studies of NRM impacts)?

If (a), what is the reason: a wrong research focus? delivery problems? If (b) why? Despite considerable IA work in crop improvement in the past, some related issues remain under-explored which deserve consideration/investigation, e.g., germplasm conservation/valuing biodiversity, passporting (value of information)/prebreeding.

Training and institutional strengthening is another under-explored area that merits serious IA attention. We know about NARS inputs (ASTI), do we know anything about outputs? Are NARS stronger today than 20 years ago? Other key candidates for IA include:

- Risk (variability) reduction, CG's contribution? Value of reduced variability.
- Improved nutrition.
- Policy change – getting some attention now.
- Priority setting - generations of models with low adoption; have we backed off too far?

The speaker challenged the group by asking what is comprehensively known about the diffusion of core technologies from major CGIAR programs. If crop improvement research is considered the major success story, even today, is it not time to update the Evenson and Gollin study which is the most comprehensive assessment to-date, although using data only up until 1997 (CIP is an exception, having updated potato varietal adoption data in 2007). It's appalling that basic data on adoption of improved varieties is not current. This should be collected on a regular and systematic basis. There was enthusiasm and general consensus that the latter topic deserves serious and priority attention. Greg Traxler organized a follow-up meeting on this subject that was attended by IAFPs representing many of the IARCs. Participants at the meeting were comfortable with the consensus view that investing in surveys of varietal diffusion in several commodities, countries, and regions were necessary to supplement and validate national and sub-national expert opinion that was the basis for the earlier work. Participants also agreed to retain the emphasis on assessing the strength of NARS in commodity improvement. Greg Traxler took responsibility to write-up the notes of the meeting together with a plan for further action. Some felt SPIA too should be involved as it was last time, although SPIA felt the motivation and leadership for this update initially should come from the IAFPs and Gates. Institutionalizing this is paramount as the database has tremendous importance as a basis for IA and research, i.e., in answering whether rates of return from CGI are declining. Other comments in the discussion related to the dearth of IA of NRM research and the need for more closely examining 'why?'

11.5 5) Institutionalizing epIA beyond IAFP specialists – presented by Tuenis van Rheenen

The issue of how epIA is organized and managed at the Centers was presented. At IFPRI, the IA unit involves the impact assessment coordinator (IAFP), senior management and researchers, with a good mix of top-down and bottom up. Ways in which IA related activities go beyond IAFP specialist include: seminars and master classes (to stay abreast of latest thinking), special working groups at Internal Program Review (building impact

indicators in projects), IA surveys (includes everyone) and commissioning IA studies. Some recommendations for building IA into the project activities both before (conducting baseline surveys, defining expected impacts with stakeholders, etc.) and after project (tracing influence, commissioning epIAs) implementation were given.

12 SESSION 12 - CONCLUDING COMMENTS FROM THE SPIA CHAIR AND THE PRESIDENT OF EMBRAPA

Derek Byerlee thanked all the IAFPs, EMBRAPA scientists, donors, special guests and SPIA members for their active participation and contribution towards making this a highly productive meeting. It is clear that there is an enormous amount and diversity of effort going on at the Centers in terms of research topics addressed, methodologies, impact pathways, etc. and this meeting has been a valuable learning and information exchange opportunity for all concerned. In addition, the meeting constituted an important step towards furthering cooperation among centers and with key partners around IA in the system; it has provided useful feedback for SPIA in terms of helping set priorities for future work; it has re-affirmed the importance and rationale for inter-center collaboration in IA; and it has emphasized the value of meta-analyses of IA (highly appreciated by donors) for rigorous and systematic learning. Presentations and deliberations will be helpful to SPIA in assessing where best it can add most value to the System given its small budget. SPIA would appear to have a continuing role to play in (i) facilitating and adding additional resources to inter-center IA initiatives, (ii) adding value with respect to methods development (iii) drawing on center IA studies to provide system wide impact estimates, and (vi) communicating IA to donors via formal (reports, briefs, AGM meetings, website) and informal means.

There is clear interest and support for the study on ESI impact; SPIA appreciates the input received during this meeting and looks forward to interacting with centers in commissioning case studies for phase II. The importance of the need for an effective communication strategy and operational plan (and rationale, objectives) was also highlighted and is well received. There is also a growing sense that epIA needs to be better integrated into learning, priority setting, and change management. The input received on PMS – and to be followed up with a special meeting to clarify some modifications in this year’s guidelines – was also helpful to SPIA. Every effort will be made to keep it simple, fair and less onerous as possible. It was agreed that a modified version of the impact indicator template and description would be circulated to IAFPs prior to submission to the ADE. It was encouraging to see the interest of a number of IAFPs in participating in one or two special symposia at the IAAE meetings in Beijing next year and yet greater efforts need to be made in reaching out to the broader evaluation community, e.g., like NONIE and 3ie (meeting in Cairo in April). As the System is in the process of a major reform it was felt to be premature to schedule the next SPIA-IAFP meeting although there is clear support for regular (biennial) meetings of this type.

Kepler Euclides Filho thanked SPIA and the CGIAR IAFPs for accepting the invitation to host their meeting at Embrapa HQ. He hoped the meeting deliberations had been fruitful and looked forward to establishing a strong collaboration with the CGIAR in impact assessment, as they already have in research.

ANNEX A: LIST OF PARTICIPANTS

CGIAR Impact Assessment Focal Point Group Meeting

EMBRAPA Headquarters, Brasilia, Brazil

November 10-11, 2008

Institution	Name	Email Address
CGIAR Centers/CPs/Systemwide Programs		
Africa Rice	Aliou Diagne	a.diagne@cgiar.org
BIOVERSITY	Elisabetta Gotor	e.gotor@cgiar.org
CIAT	Douglas Pachico	d.pachico@cgiar.org
CIFOR	David Raitzer	d.raitzer@cgiar.org
CIMMYT	Roberto La Rovere	r.larovere@cgiar.org
CIP	Guy Hareau	g.hareau@cgiar.org
ICARDA	Kamel Shideed	k.shideed@cgiar.org
ICRISAT	Jupiter Ndjeunga	n.jupiter@cgiar.org
IFPRI	Teunis Van Rheenen	t.vanrheenen@cgiar.org
IITA	Victor Manyong	v.manyong@cgiar.org
ILRI	Nancy Johnson	n.johnson@cgiar.org
IRRI	Debbie Templeton	templeton@aciarcg.gov.au
World Agroforestry	Steve Franzel	s.franzel@cgiar.org
CP HarvestPlus	Patrick Eozenou	p.eozenou@cgiar.org
CP SSA	Adewale Adekunle	aadekunle@fara-africa.org
ILAC	Jamie Watts	j.watts@cgiar.org
PURGE	John Dixon	j.dixon@cgiar.org
SC/SPIA/SC Secretariat		
SC/SPIA	Derek Byerlee	dbyerlee@gmail.com
SC/SPIA	Mywish Maredia	maredia@msu.edu
SC Secretariat/SPIA	Tim Kelley	timothy.kelley@fao.org
SC Secretariat	Sirkka Immonen	sirkka.immonen@fao.org
SC Secretariat	Ruben Echeverria	ruben.echeverria@fao.org
DONORS/Embrapa		
ACIAR	Debbie Templeton	templeton@aciarcg.gov.au

Bill & Melinda Gates Foundation	Greg Traxler	Greg.Traxler@gatesfoundation.org
Oxfam GB	James Stevenson	JStevenson@oxfam.org.uk
Embrapa	Flavio Avila	Flavio.Avila@embrapa.br
CONSULTANTS/ IA EXPERTS		
Consultant	Tom Walker	walkerts@msu.edu
USDA	Paul Heisey	pheisey@ers.usda.gov
Royal Melbourne Inst of Technology	Patricia Rogers	Patricia.Rogers@rmit.edu.au
Nat Centre for Agr Econ & Policy Res (NCAP), ICAR	P.K. Joshi	pKjoshi@ncap.res.in

ANNEX B: PROPOSED INDICATOR ON ACTUAL IMPACT

Draft Note prepared by SPIA for discussion

at the IA Focal Point Meeting, Brasilia, Nov 10-11, 2008

1. Rationale: To respond to demands by donors/stakeholders

In the context of the PMS, donors/stakeholders have been demanding an indicator that measures actual impact of a center; an indicator that will allow them to assess (at the center-level) what has been achieved in terms of CGIAR goals of poverty alleviation, environmental sustainability and food security.

2. Purpose of this note is to:

1. Present three alternate ideas to address donor demand for an indicator of actual impact
2. Present the challenges associated with meta-analysis (one of the ideas put forward)
3. Propose some guidelines and identify issues for consideration in developing guidelines for conducting meta-analysis (as an input in to discussions at the IAFP meeting); and
4. Seek input from IAFPs on the feasibility of conducting periodic meta-analysis as a way to meet the donor demand or pursue other alternatives.

3. Proposed ideas to address this need

Proposal #1: Periodic meta-analysis at the center level

One of the proposed ideas is to request each center to conduct occasional meta-analysis (e.g., once every five years corresponding with their EPMRs) of their cumulative impact stories in an attempt to compile and quantify the size and nature of the economic and non-economic impacts to-date. The approach suggested under this proposal is based on the methodology used by two recently completed SPIA-sponsored meta-analyses.¹

Under this proposal, each center will maintain a time series of documented impacts (could include outcomes also—e.g., adoption area) for (quantitative and qualitative) indicators a, b, \dots, n , (at some level of analysis—project, program, region) that are periodically aggregated across the center's portfolio and reported in agreed upon units (e.g., NPV of gross benefits attributed to center's research, number of hectares saved as a result of new technology, number of people who experienced increased income, characteristics of adopters of center's technologies, etc.). Not all centers will be required to report on actual impacts for all the indicators—but only those relevant to their research mandate and have been documented. The idea is to identify more generalized patterns of impacts from ePIA studies such that it leads to a cumulative picture of the landscape (i.e., size, scale, scope) of impacts up to a given time period that are most relevant in measuring the achievements of CGIAR goals of poverty alleviation, environmental sustainability and food security.

One of the advantages of meta-analysis is that it will provide measures of "impact" (as measured by quantitative and qualitative indicators) and can be presented as uniform

¹ See Raitzer and Kelley (2008) and Maredia and Raitzer (2006) for the practical application of this methodology, its value and limitations.

indicators relevant to CGIAR goals. In this sense it comes closest to what donors/stakeholders want. However, this idea suffers from many theoretical and practical limitations that pose several challenges (especially, the inability to rate the centers on a scale of 0-10).

The Theory and Practice of Proposed Meta-analysis

Let’s assume that there exist, for each of the CGIAR goals, indicators a, b, \dots, n that are observable and measurable, such that the observed level and direction of change in these indicators will allow us to make an assessment of the achievement (or non-achievement) of the CGIAR goals (as a result of research investments). Also, let’s assume that the measured changes in impact indicators over time, space and groups of people are additive in nature, making it possible to aggregate them across individual epIA case studies that are conducted by centers over time, space and communities.

Let’s also assume that resources are unlimited, data are available and there are no methodological constraints in measuring and attributing impacts. Then, in theory, it is possible to construct a time series of values of impact indicators a to n that can be updated every x number of years. Under this perfect condition, information on realized impacts (positive and negative) is available for all impact indicators and it is possible to get a picture of the landscape of impacts up to time period T , as shown in Figure 1.

Year	Impact Indicators			
	a	b	...	n
1	a_1	b_1	+/-	n_1
2	a_2	b_2	+/-	n_2
3	a_3	b_3	+/-	n_3
...	a_{\dots}	b_{\dots}	+/-	n_{\dots}
T	a_T	b_T	+/-	n_T
Total	$A = \sum_{t=1}^T a_t$	$B = \sum_{t=1}^T b_t$		$N = \sum_{t=1}^T n_t$

Figure 1 Impact matrix for potential impact indicators a, b, \dots, n for time period $1,2,\dots,T$

In Figure 1, the observed values of impact indicators a,b,\dots,n for any given year represent aggregate impacts across research portfolio ($p = 1, 2, \dots,P$) and geographic locations ($l=1,2,\dots,L$), sourced from individual epIA case studies conducted by the center up to time period T . Thus,

$$(1) \quad a_t = \sum_{p=1}^P \sum_{l=1}^L a_{plt} .$$

Thus, the total estimated impacts A, B,\dots, N are center-level indicators of impacts aggregated across programs/projects, locations and time period (from several different epIA case studies).

This theoretical possibility of conducting meta-analysis, however, suffers from the following limitations:

1. Due to the fact that a research output generates impacts over time and space at varying rates, the estimated values of total impacts A, B, \dots, N (as given in Figure 1) will not capture all the potential impacts of research investments up to time T ?. The estimated stream of impact indicator series A, B, \dots, N will have to be viewed as a cumulative stock of impacts up to time period T , rather than an indicator of annual (or x number of years of) impact corresponding to research investments in year t (or years $t, t-1, \dots, t-x$).
2. Due to the fact that research is not a homogenous activity and target research domains vary across research programs/center, there are no set targets (or standards) for what constitutes an acceptable, good, better or best “impact achievement” over time, space and communities. The estimated stock of impacts at any time t for a research program (or a center) will thus not be comparable across programs (or centers) because of the lack of such a standard on what constitutes an acceptable or unacceptable “impact.”²
3. The documented impacts would be realized as a result of several different outcomes and contributions from several different players. Attributing those impacts to the specific center may not always be feasible. Hence, it may not be possible to compare the aggregate impacts from the meta-analysis with the cumulative investments of the center up to time T .

In practice, the following conditions add to the list of constraints that will make the meta-analysis differ from the ideal theoretical possibility depicted in Figure 1: First, neither resources are unlimited nor data easily available to assess/measure the “level and direction of change” in all the impact indicators. Second, methodologies are not developed uniformly to measure changes in all potential types of impact indicators a to n (assuming that there exist observable indicators for all CGIAR goals) to make an assessment of the achievement (or non-achievement) of the CGIAR goals as a result of research investments. Due to these realities, centers conduct impact assessment for selected projects and programs (i.e., do not cover their entire research portfolio $p = 1, 2, \dots, P$). As such, the documented impacts by individual epIA case studies are not continuous over time and space and the assessed impacts are not comprehensive in measuring the ‘level and direction of change’ in all the potential impact indicators. Third, economic indicators embodied in the concept of consumer and producer surplus are by far the most heavily assessed indicators in impact studies for agricultural R&D. Centers and/or research portfolio that generate outputs that are less amenable to measuring economic impacts (e.g., policy, NRM, capacity building, etc.) are therefore disadvantaged. Thus, due to the above realities and conditions:

At any given time period t , the time series of changes in impact indicators actually assessed/measured will be patchy, incomplete and not up-to-date.

Because of the lag in measuring impacts (either due to lags in realization of impacts or methodological constraints) one will never be sure whether the estimated values represent the actual impact of the center to date or not.

Changes in the estimated values over time may not be reflective of productivity changes of a

² Although, in theory it is possible to compare each indicator a, b, \dots, n to get relative ranks after controlling for factors such as the size and age of research centers.

research center; it could merely be a reflection of new evidence of documented impacts made available over time as a result of methodological innovations and data availability.

Due to limitation #2 noted above, it is not possible to construct an “impact indicator” on a set scale of 0-10 for across-center comparison purpose—something donor/stakeholders may want for their decision making process (and one of the reasons for the creation of the PMS).

To cope with these imperfections of the real world, we present the following ideas on what the indicator of “actual impact” and the process involved using the approach of meta-analysis will look like.

The Proposed Meta-analysis Approach—Practical Application

Box 1 describes the three step process envisioned in constructing indicator(s) of actual impact through meta-analysis. As a first step, this approach requires forethought and planning of epIAs by keeping in mind the data and information needs for the subsequent meta-analysis. The second step involves collecting data/information from each of the completed studies (that meet the quality criteria) in a standard format (perhaps a spreadsheet). The third step is then to conduct the meta-analysis based on the meta-data collected from each epIA.

Box 1: Proposed steps in conducting meta-analysis

Step 1: Plan and implement epIAs keeping in mind the ease of integration of the study into a subsequent meta-analysis (see steps 2 and 3)

Step 2: Collect data and information from each of the completed epIA case study (which will form the data source for step 3--meta-analysis)

Example:

Ex-post Impact Study (full citation): _____

Subjected to external peer review? Yes/No

Subjective assessment on whether the study meets the good practice guidelines on transparency (clearly derived key assumptions, attribution of data sources, and explanation of data treatment)? Yes/no (or rating on a standard scale)

Subjective assessment on whether the study meets the good practice guidelines on analytical rigor (appropriate data treatment, use of representative data, plausible counterfactual scenario developed, plausible institutional attribution)? Yes/no (or rating on a standard scale)

List/identify outputs (attributed to the center’s research) being assessed by this study: _____

Evidence of outcomes and impacts documented in this epIA (complete the following table, as applicable)

Time period	Outcome indicators			Indicators of documented impacts				Qualitative assessment of documented impacts	
	Examples			Examples				Examples	
	Adoption area	Evidence of uptake of recommended practices	Evidence of new or improved govt. policy	Economic value of estimated benefits (GARB)	Number of hectares saved from deforestation	Number of people averted from hunger/malnutrition	Number of people lifted out of poverty	Profile of people impacted (e.g., small scale producers, urban consumers, mother and children, women)	Profile of geographic regions impacted (low potential, high potential, political boundaries, rainfed/irrigated)

1									
2									
...									
T									
Cumulative over time (if feasible)									

Step 3: Conducting the meta-analysis

Data source: List/identify epIA studies over the past x number of years that meet some minimum threshold criteria for quality and rigor (or group them into 2-3 categories based on quality and rigor). These will form the pool of impact studies for meta-analysis

Develop a database/spreadsheet table for each indicator of outcome and impact with the source of data (i.e., epIA studies) as columns and time period of documented evidence as rows

If quantitative evidence is reported (e.g., adoption area, economic benefits, number of people) ensure consistency in units being reported across studies.

Exclude duplicative evidence of the same outcomes and impacts documented and reported in multiple studies (retain the data from the highest rated study)

Aggregate the quantitative data for a given indicator across epIAs

Reporting the results: i) For indicators measured in monetary units, present the aggregate and cumulative total as NPV (using an appropriate discount factor and base year—which may be standard across all the centers); ii) For indicators that are additive over space but not over time (e.g., adoption area), present the time series of aggregate values across epIAs; iii) Analyze quantitative data (e.g., monetary benefits, adoption area) by qualitative information related to outcomes and impacts (e.g., profile of adopters, type of technology, geographic regions, types of beneficiaries, etc.) to present the size, scale and scope of the documented impacts up to time t; iv) If a study included in the meta-analysis pool is completely qualitative, present the documented evidence as a case study of an impact story.

Issues for Further Discussion Regarding Proposal #1

Time frame: Centers will conduct meta-analysis every five years corresponding with their EPMRs.³ The first meta-analysis to be conducted will cover, compile and present a cumulative picture of the size and nature of the economic and non-economic impacts to-date (quantitative + qualitative). Meta-analyses conducted thereafter will be an update over the previous one.

Tasks for the center: Each center will maintain a time series of documented impacts and outcomes for indicators related to CGIAR goals based on impact studies conducted at the project, program, or geographic region level. These documented impacts/outcomes will be aggregated across the center's portfolio and reported in agreed upon units (e.g., NPV of gross benefits attributed to center's research, number of hectares saved as a result of new technology, number of people who experienced increased income, etc.) every five years.

Guidelines for conducting meta-analysis: The meta-analysis (i.e., reporting of cumulative aggregate values of impact indicators) will be based on guidelines to be developed by SPIA (or jointly by SPIA and center IAFPs) to maintain quality/rigor, credibility and uniformity of results across centers. The following discussion includes a starter list of issues that will have to be addressed in developing the guidelines and need further discussion:

Acceptable types of ePIAs on which the meta-analysis is based: For example, studies must:

- Be a peer reviewed publication
- Must document impacts of outputs/outcomes that can be traced to the center's research efforts (in a plausible manner)
- Must be based on sound methodology and credible sources of data (must pass the standards for transparency and rigor).

Other...

Also, clear guidelines need to be developed on how and who determines the acceptability of a study for meta-analysis (for example, who makes the subjective assessments in steps 2.c and 2.d noted in Box 1?).

Types of indicators to be reported: What are the indicators of impacts that should be included in the meta-analysis that are: additive across time, space and communities; are meaningful in conveying the size/scale of impact as related to CGIAR goals; and potentially cover all (or major) types of research conducted by CGIAR? (Note however, that not all centers will be expected to report on all the indicators; a given center may report on a sub-set of indicators relevant to their research mandate and documented in the past).

Rules of aggregation: For example, rules pertaining to discount rate and currency to be used (if impacts are reported in monetary terms) in meta-analysis, rules of no double counting (i.e., if the benefits of a center's research output for a given time period and location are estimated/documentated by multiple studies they should be included only once in the meta-analysis, assigning weights (if any) in aggregating costs/benefits across time, space and communities, rules for presenting qualitative data/information on aggregate impact

³ This is open for discussion. The other alternative would be to conduct the meta-analysis every five years across all the centers (i.e., at the same time).

indicators, etc.

Rules of attribution: How to treat the contribution of others in realizing the documented impacts? Should the meta-analysis be based on study-defined attributions of impacts to a center's research (where such information is available) or only report total impacts for the sake of uniformity?

Treatment of future/projected impacts: Most epIAs also include estimates of impacts in to future (i.e., ex ante assessments) based on some assumptions. Guidelines will need to be developed on how to treat these ex ante estimates of impacts in the meta-analysis. Should the meta-analysis be strictly based on the ex-post period defined by the study or by the time frame of the meta-analysis?

Rules for periodic update of meta-analysis results: Will the meta-analysis to be conducted every five years be a cumulative analysis of impacts up to that year (allowing for updates of evidence of impacts prior to the past five years) or strictly an analysis of additional impacts realized in the last five years?

Presentation and reporting of results of meta-analysis: Should the results be published by each center as a Report that is published every five years (based on their calendar of EPMRs) in a format that is uniform across centers and becomes one of the background documents for the EPMR? Or SPIA prepare a report every five years based on the meta-analyses that are completed by each center up to that time? Does SPIA play an "evaluator/reviewer" role as it does for the PMS or only as a resource/facilitator in this exercise?⁴

Proposal #2: Periodic impact assessment of Centers based on in-depth analysis of "impact claims"—the case study approach

The idea here is to employ "impact pathway analysis" with the pathway spanning from research output, outcomes and impacts, and against which each center will provide verifiable impact claims periodically (say every five years corresponding with EPMRs).⁵ Each center will be asked to submit up to x number (to be determined based on the age and size of the center) of case study examples of impact stories that contextualizes their research along the impact pathway, allows them to demonstrate how impact of their research is achieved and substantiated, and the breadth and depth of those impacts. These case studies of impact stories will be submitted by the centers based on a content guideline, which will be carefully drafted to allow centers to present appropriate evidence to demonstrate the depth of impact their research has achieved in a variety of domains ranging from economic, environmental and social impacts.

Rules can be set on the number of case studies, how they are selected and the time frame of impact claims made. The methodology for making an assessment of the achievement of impacts (i.e., an assessment of whether a center's research is creating significant and substantial public value) will involve an expert panel applying judgement to a combination

⁴ Note that due to the limitations noted earlier, it will not be possible for SPIA to rank/rate the indicators of 'actual impact' on an absolute scale of 0-10. But it can comment on the quality and rigor of the method used in the meta-analysis by different centers.

⁵ This is a similar approach to how the outcome indicator is currently assessed in the PMS—but on an annual basis.

of qualitative and quantitative indicators that support the claims.

Pros and cons:

The case study approach of in-depth analysis of impact claims for a given line of research (or a project) based on verifiable evidences spanning research outputs, outcomes and impacts does not have the main advantage that the meta-analysis has—i.e., it does not result in a value of actual impact of a research center. However, it does provide an opportunity for centers to demonstrate the value of their science—both in terms of excellence and relevance, by taking an in-depth look at selected research projects/themes every y number of years. This approach also allows for heterogeneity in types of research included in the assessment (as it is not solely based on quantitative indicators). Unlike meta-analysis, this approach can result in a value “rating” that can be presented on a scale comparable over time as well as across centers. In this sense, this approach comes closest to a PMS system (albeit at a frequency of 5 years), if donors/stakeholders do want an indicator for comparison purpose.

Proposal #3: System-level meta-analysis on an annual basis

Another idea proposed is to report on impacts for the CGIAR system based on the EPMRs (“impact nuggets”) and on impact studies finalized during the year. This will be an attempt to capture/measure “system level” impacts on an annual basis.

This proposal is very similar to proposal #1, except that the meta-analysis will be done at a system level on an annual basis. Guidelines will be developed for centers to conduct steps 1 and 2 as described in Box 1. An appropriate entity (SPIA?) at the system level will then be in-charge of doing step 3.

Pros and cons:

The idea of conducting system-level meta-analysis on an annual basis based on EPMRs and impact studies finalized during the year is an attractive one as it brings the focus on system-level achievements and impacts, and eliminates the pressure of comparison across center. But are donors willing to take the pressure off each center for accountability and take the macro-approach of the overall system-level impacts to make investment decisions? Also, who will be the appropriate entity to conduct such system-level meta-analysis on an annual basis? What are the time and resource needs for such an exercise? These and the issues identified on pages 5-6 need further discussion.



CGIAR

Science Council Secretariat
c/o FAO
Viale delle Terme di Caracalla snc
00153 Rome, Italy

t +39 06 57056782
f +39 06 57053298
e sc-secretariat@fao.org

www.sciencecouncil.cgiar.org