

April 29, 2004

Dear Ian and Francisco:

We are writing to brief you on the progress made by the Working Group on Performance Measurement (WGPM) the ExCo appointed last year. We were pleased that the recommendations made by the Working Group at the end of Phase 1 were endorsed by ExCo and the CGIAR. With that "green light", we began Phase 2 last month and have first developed a plan for designing, developing and implementing a performance measurement system for the CGIAR. This is detailed in the attached draft paper prepared by the WG, with the assistance of our technical experts and resource persons (TERPS). This is very much a working draft and is being reviewed by the WG members as we are forwarding it to you for information. It is being shared in this draft form so that the WG can benefit from the inputs of various partner groups who will be meeting over the next two weeks and have included the work of the WG in their agendas (such as the Science Council, CBC, CDC, and ExCo).

We would like to share with you the WG's thinking on the key principles that should guide the design and development of the PM system. These were discussed at length at our meeting on April 22, 2004, and we would welcome any reactions, especially from the our colleagues representing the donor community:

1. The Working Group agreed at its April meeting that the implementation of this system should proceed as quickly as possible, but that we should not sacrifice quality for speed.
2. The PM system should be based largely on self-assessment by the centers, utilizing available information to the maximum extent possible. A verification system should also be built, relying on the Science Council, the Boards, the Internal Audit Unit, and the EPMRs.
3. The Working Group is aware of the many concerns about use of such a system for resource allocation. In the WG's view, a PM system is needed, first and foremost, as a management tool and a tool for learning by those who have the responsibility to manage the centers. The same information, in whole or in part, could be used by all stakeholders for other purposes. The PM data can be an input to decision making by the donors, provided it is not the only information on which decisions are made and sufficient care is exercised in integrating and assigning weights to the various indicators.

4. We were briefed by our new SC Chair and WG member, Per Pinstrup-Andersen, about the Science Council's intention to revamp the CGIAR's Medium-Term Planning process. This is a welcome development from a PM standpoint because center MTPs contain some of the key data needed to measure results of center activities. The planned improvements in the MTP process would help generate more accurate information on program performance, along with the judgments of the Science Council.
5. We had considerable discussion of timing. The Science Council is concerned that it may not have milestones in place across the system for 2004, and that without these, the 2004 data may not serve as an effective pilot—at least for some of the proposed elements. Nonetheless, there was general agreement that centers should begin to build data systems as soon as the indicators are chosen. For sake of transparency, the raw data reported by the Centers should be made public, even if some donors decide not to make their analysis of these data public.

We also discussed the set of performance indicators used by the World Bank for allocating a small part (12.5 percent) of its 2004 contributions to the System. These indicators were chosen, developed, and reported by the Center Directors. Three performance areas were included:

- Producing and disseminating knowledge as public good
- Capacity building and partnership for helping others produce public goods
- Financial health of the center

The WB exercise serves in many ways as an informal pilot for the current effort. We have been informed that the WB would shift to the fuller set of indicators emerging from the WGPM once these are ready, instead of continuing to use the limited set of indicators used for 2004.

We welcome ExCo's reactions to the plan developed by our WG.

With best wishes,

Kevin Cleaver and Luis Arango Nieto  
Co-Chairs

# **A Plan for Designing, Developing, and Implementing a Performance Measurement System for the CGIAR**

*Draft2*

*April 27, 2004*



CGIAR Working Group on Performance Measurement

## Table of Contents

1. Introduction.....	3
1.1 The Worldwide Move Towards Performance Measurement.....	3
1.2 Why Develop a Performance Measurement System for the CGIAR?.....	3
1.3 The Process to Date.....	4
2. Performance Measurement in the CGIAR.....	5
2.1 Implications for CGIAR of Advances in Performance Measurement .....	5
2.2 The Purposes of Performance Measurement in the CGIAR.....	5
2.3 Principles for Performance Measurement in the CGIAR .....	7
2.4 Elements of a Performance Measurement System for the CGIAR.....	8
2.3 How will the Performance Measures fit into the current CGIAR Planning and Evaluation Processes? .....	10
3. Plan for implementation .....	14
3.1 Indicator Development and Choice.....	14
3.2 Communication with Centers through Focal Points .....	14
3.3 Data collection and reporting procedures .....	15
3.4 Timeline .....	16
Annex 1: Composition of the CGIAR Working Group on Performance Measurement (WGPM) .....	17
Annex 2: Instructions to Specialist Teams in Choosing and Developing Indicators .....	18
Annex 3: Description of Elements .....	19

# 1. Introduction

## 1.1 *The Worldwide Move Towards Performance Measurement*

National and state governments around the world are implementing requirements for performance measurement and annual performance reporting by government agencies. This trend was fuelled by the economic downturns and increasing government deficits of the 1980's and 90's and the concomitant decrease in public confidence. The use of information on the results of government programs was seen as a rational basis for encouraging the efficient use of resources and making the hard allocation decisions that needed to be made. Use of performance information as a tool for greater transparency and reporting of government accomplishments has been seen as a response to lowered confidence.

Performance measurement is also a response to management's need for current information in a world of increasingly fast paced change. In the early 1970's Canada and the U.S. had implemented policies requiring agencies to assess how well their programs were working through detailed analytic studies, called "evaluations" that were conducted from time to time. Eventually, dissatisfaction grew with these studies because of the time they took to complete and because the costs of conducting them meant that only a limited number of programs could be studied each year. Consequently, managers began developing "performance measurement" or "performance monitoring" approaches that collected and reported performance information on a recurring and timely basis. These approaches provide an "early warning" system that allows managers to detect and solve potential problems at an early stage.

Research organizations that are government sponsored or that obtain a significant proportion of their funding from government sources have been affected by the requirements for performance measurement, and many have moved to implement performance measurement systems. Some examples are discussed more fully in the *Sourcebook on Performance Measurement in Research Institutions and Programs*.<sup>1</sup>

## 1.2. *Why Develop a Performance Measurement System for the CGIAR?*

In recent years, the need for taking a fresh look at performance measurement issues in the CGIAR has been stressed by various CGIAR bodies such as CGIAR Executive Council (ExCo, Sept. 2002) and the Working Group on the Establishment of the Science Council (WGSC, August 2002). Also, the CDC expressed interest in developing measures of Center performance, as part of an overall effort to reform the CGIAR's evaluation system.

The CGIAR Working Group on the Science Council urged a "much wider use of self-assessment in reviewing and enhancing the quality of science in the CG System" and suggested consideration of a performance assessment framework based on quantitative indicators. The recent iSC proposal on "Changing Monitoring and Evaluation in the CGIAR System" echoed the sentiment by pointing to the need for a continuous self-

---

<sup>1</sup> CGIAR Secretariat, Performance Measurement of Research Institutions and Research Programs - A Sourcebook, Draft, Washington, June 2003.

assessment that covers such matters as the relevance of research and related activities, quality of science, outputs and impacts, the efficiency and effectiveness of operations, including partnerships (iSC Secretariat, June 2003).

At its third meeting ExCo concluded that the System Office should bring recommendations to ExCo on how to approach the question of performance measurement on a system level. As a follow up, the System Office developed a concept note on establishing a Working Group on Performance Measurement in the CGIAR. A two-phase approach was recommended: in Phase 1 the WG would develop options for performance measurement and in Phase 2 the chosen option would be designed, developed and implemented.

### **1.3 The Process to Date**

The System Office established the Working Group on Performance Measurement in May, 2003, under the Co-Chairmanship of two ExCo members: Kevin Cleaver (ExCo/FC Chair) and Luis Arango (ExCo/PC member). Other members represent a cross section of interests and expertise from inside and outside the CGIAR (see Annex 1 for the membership of the WG). Three sets of activities were carried out in preparation for the first meeting of the WG on September 5, 2003:

- (1) The CGIAR Secretariat prepared a *Sourcebook on Performance Measurement in Research Institutions and Programs* as background on approaches and methods of performance measurement being used in similar organizations globally.
- (2) Members of the WG shared additional information relevant to the objectives of the exercise (e.g., papers, articles);
- (3) A sub-group of the WG (made up of technical experts and resource persons<sup>2</sup>) met for a two day workshop on August 11-12, 2003 to discuss and outline performance measurement options that could be considered by the WG at its September 5, 2003, meeting, as a means of facilitating the task of the WG.

Since the Working Group issued its recommendations, several more steps toward the PM system have been taken:

- (1) The report was shared with ExCo and the CGIAR and comments were sought until December 31.
- (2) In January 2004, the CGIAR Director reported to ExCo that “the CGIAR has endorsed the Working Group's recommendation to proceed with the development of a performance measurement system along the lines suggested in the report. The suggestions made by the members will be taken into account by the WG during the design phase. As agreed by the CGIAR, ExCo would monitor the development of the performance measurement system.”

---

<sup>2</sup> Flavio Avila, Stan Divorski, Ruben Echeverria, Doug Horton, Maria Iskandarani, Mortimer Neufville, Selçuk Özgediz, Ray Rist.

- (3) The technical subgroup<sup>3</sup> met again on March 30-31, 2004, to review the comments received and develop a draft for design and development of the performance measurement system.
- (4) The Working Group met on April 22, 2004, to discuss the proposed plan and agree on next steps.

In this document, we put forward a plan for Phase 2 of the performance measurement project. Chapter 2 describes potential purposes and uses of performance measurement, lays out principles for the performance measurement system, introduces the ten performance elements the WG recommends, and outlines how such a system could fit into the planning and evaluation processes of the CGIAR. Chapter Three presents an implementation plan and timeline.

## **2. Performance Measurement in the CGIAR**

### **2.1 Implications for CGIAR of Advances in Performance Measurement**

Common conceptual models used for monitoring organizational performance show that performance measurement should be more than a simple measurement of a set of output indicators. Generally, the CGIAR Centers use *resources and inputs* (funds, personnel, equipment, and facilities) to undertake their *research operations* in order to produce *outputs* (e.g., management practices and research methodologies) for the benefit of farmers and other users. The *outcomes and impacts* of adopting or applying these outputs are measured by their effects on, for example, production cost, yield, income, natural resources, etc. Ideally a performance measurement system would capture this whole chain of causality.

Experiences of others also show that for performance measurement to succeed there must be agreement on the purposes to be achieved by the performance measurement system. Another lesson is that a PM system should not be “copied” from another organization, rather it needs creative development, stepwise adjustment, and, most importantly, commitment from all stakeholders. Moreover, the CGIAR has its own history, established values and other special characteristics that need to be taken into account in designing such a system. Therefore, one should not be surprised that development of such a system would take time, and start-up problems would need to be overcome before a functioning PM system could deliver robust information for use by all.

### **2.2 The Purposes of Performance Measurement in the CGIAR**

The CGIAR has a strong and well-recognized track record when it comes to performance. The centers are known as “centers of excellence” because of the reputation of their scientific achievements. External reviews of the System and of the centers have documented the significant achievements and impact of the centers.

The CGIAR and the centers also employ a range of tools for monitoring the quality and impact of the System’s activities. However, the existing tool set needs to be adapted to

---

<sup>3</sup> Flavio Avila, Susan Cozzens, Ruben Echeverria, Doug Horton, Sirkka Immonen, Maria Iskandarani, Mortimer Neufville, Selçuk Özgediz.

the needs of a rapidly changing environment and the demands from stakeholders. The various partners and stakeholders in CGIAR have been asking a variety of questions about the continued financial and institutional health of the centers, the performance of the Boards, maintenance of the centers' comparative advantage in quality of science, the range of outputs generated annually, etc.

A well-designed performance measurement system could provide answers to many of these questions. The information generated through such a system would have a variety of potential uses, including:

- serving as a tool for decision-making and/or performance management by the centers,
- demonstrating accountability,
- benchmarking, and
- aiding in resource allocation.

***Decision-making and Performance Management.*** Centers could benefit from having a performance measurement system that not only monitors accomplishments, but also monitors and provides timely information on their potential to perform well in the future. Such information, appropriately adjusted, could be used by the centers as part of their internal *performance management* effort. It would help them to identify areas needing early attention and support efforts to learn from their own performance, innovate, and make appropriate changes to improve performance. In addition, having uniform indicators of performance across the centers would enable the System to monitor its progress towards desired goals—as a System.

***Accountability.*** CGIAR's structure as an alliance of equal and autonomous partners (i.e., centers) is similar to structural models that many governments have been experimenting with, and performance measurement in such a system is important for many of the same reasons. Each must do its part if agricultural research is to contribute to reducing hunger and poverty; and each is accountable to the other for the results they achieve, how it achieves results, learns from experience, uses resources efficiently, and improves its contribution to the common goal. A performance measurement system is a valuable tool in providing the information centers need to demonstrate these accountabilities to each other, to the stakeholders, and to their investors.

***Benchmarking.*** Performance measures that are consistently and uniformly applied across centers support institutional learning and change by allowing center performance to compare themselves against other centers, and the collective CGIAR performance to be benchmarked against similar organizations. Care is of course needed in making comparisons. Benchmarks would allow a center, or CGIAR as an institution, to identify areas of high performance so that best practices can be identified and adapted, as appropriate, to the needs of individual centers.

***Resource Allocation.*** In the public sector, one frequently cited purpose of performance measurement has been to use the information as a basis for resource allocation. Unlike the public sector, in the CGIAR there is not one, but 63 (as of this writing) actors which

face a resource allocation problem, mostly on a yearly basis. A well-conceived performance measurement system could provide valuable information to each of these “investors” that they could take into account when making their resource allocation decisions. This could be but one factor that enters into each investor’s calculus in allocating resources. Because the reasons for low performance one year can be complex and varied and can include factors largely outside the control of the organization, the performance data need to be used cautiously. When the reasons for low performance one year are assessed, the appropriate response could very well be to increase, rather than decrease resources.

The WG agreed that the CGIAR’s PM system should serve multiple purposes. The primary purposes should be the promotion of high Center performance and accountability towards achieving goals. The system should be developed and implemented in full consultation with the Centers to encourage Center buy-in and use of the system.

The Working Group sees the establishment of a PM system secondarily as a tool for improving transparency and demonstrating accountability and third as an additional instrument that could be used by interested members in making their resource allocation decisions. However, the WG stresses that performance measurement information should be used thoughtfully in assessing and comparing Center performance. It is vital that decisions made on the basis of assessments of Center performance be made with a full understanding of the relevant circumstances. Performance information is not a substitute for detailed evaluations and reviews. Rather it should be viewed and developed as a complement to existing processes for assessing Center performance.

### **2.3 Principles for Performance Measurement in the CGIAR**

The WG group proposes that the performance measurement system initially rely, to the extent possible, on center self-reporting, which would require a process for external validation of the reported information. As the system evolves, the need for additional information collection, such as regular client surveys may become evident. Even where there may be existing data to draw on, Center information systems may need to be upgraded and some aspects of Center information collection harmonized to develop performance information that meets the needs of the CGIAR.

The Working Group has had much discussion of the need for a balanced set of indicators and measures. Reliance on individual indicators could divert Center performance in a single direction, to the detriment of achieving other organizational goals. There will also need to be a careful balancing of quantitative and qualitative measures. For example, some center accomplishments reported qualitatively may be of such significance that they could outweigh performance on other indicators. These would not be made visible if there was exclusive reliance on quantitative measures.

Also, because of the diverse nature of Center missions and mandates, variations in their size, and the different circumstances under which each operates, what counts as good performance on an indicator may vary from center to center. Indicators should be chosen to apply to as broad a range of centers as possible. For some indicators, a minimum acceptable level of performance may be defined. For many, centers will need to set their own goals for improvement based on consultation with the System Office, the Science

Council, and the center's own board. For outcomes, the PM system should use the milestones identified in the center's medium term plan.

The WG identified eight elements of performance, which are summarized in the next section. The discussion at the WG meeting revealed issues with all of them that need to be addressed in the next stages of development. There was consensus that the results elements would be the most challenging, because of the inherent difficulties of assessing the outcomes and impacts of scientific activity. There was much concern that although information on Center outputs could be collected annually, this would not be true of outcomes and impacts. Impacts were considered especially challenging to address via performance measures. Because of these challenges, the Science Council should play an important role in the design of the PM system, particularly with regard to designing indicators of science quality, relevance and results.

#### **2.4. Elements of a Performance Measurement System for the CGIAR**

The conceptual models of performance reviewed in the September, 2003, report of the Working Group are not specific to international agricultural research. They must be adapted to the unique nature of CGIAR center accountabilities and international relationships, and to the specific and varied nature of their research. For this reason, a performance measurement system for the CGIAR should be developed from the ground up, not by "copying" the model used in another organization. It is also clear that the generic models presented do offer many useful insights into what factors contribute to performance in organizations.

The experiences of others, and the literature reviewed, suggest that two groups of factors should be thought of as candidates for a performance measurement system:

1. Elements reflecting the organization's *potential to perform*;
2. Elements reflecting the *results* dimension of the organization's work, i.e., its outputs, outcomes, and impacts;

The Working Group recommends a set of eight elements that fit the special circumstances of the CGIAR for inclusion in the CGIAR's performance measurement system. These include the following:

##### **Results**

- Outputs
- Outcomes
- Impacts
- Stakeholder Perceptions

##### **Potential to Perform**

- Quality and Relevance of Staff
- Quality and Relevance of Programs
- Governance and Institutional Health
- Financial Health

The following is a brief presentation of each of these eight elements. Examples of indicators that could be developed for each of these elements are presented in Annex 3.

## Results

### **ELEMENT 1: Outputs**

At the core of assessing “how well a center is doing” is information on the extent to which the center provides outputs that lead to outcomes and impacts consistent with its mission and objectives. The CGIAR Logframe defines “outputs” as the “products (tangible/intangible) delivered by projects.”<sup>4</sup> Examples of outputs are germplasm and germplasm improvement techniques; procedures for germplasm conservation; management practices and research methodologies; and improved policy analyses and techniques for policy formulation. The types of outputs that are relevant for a particular center’s work depend on its mission and programs.

### **ELEMENT 2: Outcomes**

“Outcomes” refer to the use of center outputs by stakeholders and clients, e.g., changes in knowledge, attitudes, policies, research capacities and agricultural practices. In the terminology of the CGIAR Logframe, they are “purposes.” Typically the extent of use of scientific research has been judged qualitatively through expert opinion provided by peer reviews. Other organizations frequently report on use of their results through examples. There has also been some development of quantitative indicators of short term outcomes, such as the volume of research stimulated elsewhere.

### **ELEMENT 3: Impacts**

“Impacts” refer to the longer range social, environmental and economic benefits that are consistent with the center’s mission and objectives, e.g. increased agricultural productivity, improved food distribution, etc. While the intended impacts of the CGIAR system are very well defined, the specific role of the centers is often difficult to assess. The desired impacts are often influenced by a number of factors, with science playing only one role. In the terminology of the Logframe, this category corresponds to “intermediate goals” (“direct benefits resulting from the uptake of innovations which include outputs from the CGIAR”).

### **ELEMENT 4: Stakeholder Perceptions**

One goal of the performance measurement system is to make objective information on center performance widely available. The perception of performance among stakeholders, however, can be equally influential. The Working Group recommends that as part of the performance measurement system, feedback from donors, partners, and users of center results be gathered on a periodic basis on elements of center performance.

### **Potential to Perform**

### **ELEMENT 5: Quality and Relevance of Staff**

The CGIAR is a “science system.” Central to its credibility is the quality and relevance of research activities ongoing in the centers. The people who work at the CGIAR centers are

---

<sup>4</sup> CGIAR Resource Allocation, December 2000, p. 23.

the key input to the effectiveness of the system. This element provides indicators of whether centers are recruiting and retaining staff capable of high quality research, as judged by international standards. The skills balance on the staff also needs to reflect the needs of each center's programs, and the staff should be moving toward balance by region and gender.

### **ELEMENT 6: Quality and Relevance of Programs**

The quality and relevance of center programs provide an indication of the extent to which current research will continue to yield accomplishments consistent with centers' missions and objectives. The standards for quality are set by the international research community, and the standards for relevance are set by the CGIAR's own goals and priorities. Each center translates system goals into its own goals and programs. The judgment of relevance for a center, then, is in relation to its own mission and goals.

### **ELEMENT 7: Governance and Institutional Health**

Continuing success and improvement depend upon a center's being effectively governed and on the health of its internal institutional climate. In a performance measurement system, it is especially important to have indicators of the extent to which the organization's culture fosters performance. Strengthening their orientation to learning and change and extending it to such new areas as greater responsiveness to stakeholder needs will help centers ensure that past levels of success will continue into the future and perhaps be exceeded. Measurement of the extent to which centers enter into partnerships and of the nature of these partnerships provides an indication of their capacity to maintain and enhance their performance under changing circumstances. It is also important to monitor the extent to which the organization has in place key processes that support governance and contribute to institutional health.

### **ELEMENT 8: Financial Health**

Financial health – the effective acquisition and management of financial resources -- is fundamental to a center's performance. For an organization to make reasonable progress towards its goals and objectives, it must have the required resources available when needed. Sufficient flexibility in its financial situation helps an organization be innovative, take advantage of new opportunities and adapt to changing circumstances. Reporting on financial health is also a fundamental aspect of any accountability relationship and helps funding organizations to make resource allocation decisions.

### **2.3. *How will the Performance Measures fit into the current CGIAR Planning and Evaluation Processes?***

Centers are currently burdened with a variety of ad hoc evaluations and data collection requirements. The PM system should be developed with these requirements in mind so that it might serve as a possible substitute for some of the information collection by individual CGIAR members and other stakeholders.

### ***Key Strategic and Operational Planning Instruments***

The CGIAR has a number of planning and evaluation instruments at different levels, which should not be seen as detached from each other, but rather build on and complement each other (see Figure 1).

The major **system-wide strategic planning and priority setting** for the CGIAR is driven through the Science Council, and the Science Council is responsible for ensuring that centers' programs are aligned to the system wide strategy. Subsequently, centers' strategies are translated into **Medium Term Plans (MTPs)**, three year rolling plans that are updated annually. The right time to compile and report on the performance measures in the new system may be when the MTP is being compiled and delivered. A center's report on outputs, outcomes, and impacts needs to be evaluated in relation to its system-endorsed MTP.

On program and project level the **Logframe approach** was introduced as a planning tool in 1998. The intention was to increase the efficiency and effectiveness of research and research management at all levels. The main objectives were: (1) to have an output oriented approach in research planning and research management (2) increased accountability at all levels, (3) increased transparency, (3) increased compatibility of management tools. Although it has not been comprehensively implemented across the system as a management tool, some centers used the Logframe and the centers continue to report their program and budget proposals using the output categories of the Logframe.

The lessons learned from Logframe implementation would be of great value for the design of a PM system. It would also be helpful if the categories used in the PM system were the same as those used in the Logframe. The concept of "outcomes" in the proposed PM system corresponds well to that of "purpose" in the Logframe. Likewise, "impacts" in the proposed PM system corresponds well to "intermediate goals" in the Logframe.

### **Monitoring and Evaluation Instruments**

The CGIAR System and centers have developed an elaborate system of monitoring and evaluation, consisting of center reviews, inter-center thematic or system-wide program reviews, and System reviews. These include the following:

- **CGIAR System reviews** – undertaken periodically (three since the founding of the CGIAR in 1971) to evaluate the overall performance of the CGIAR System at large;
- **External Program and Management Reviews of Centers (EPMRs)** – organized by the Science Council and the CGIAR Secretariat every five years to evaluate respective Centers' programs and management;
- **Centre Commissioned External Reviews (CCERs)** - commissioned and managed by the Centers themselves, as in-depth evaluations of the relevance and quality of science in specific research programs or as assessments of specific areas of operations and management;

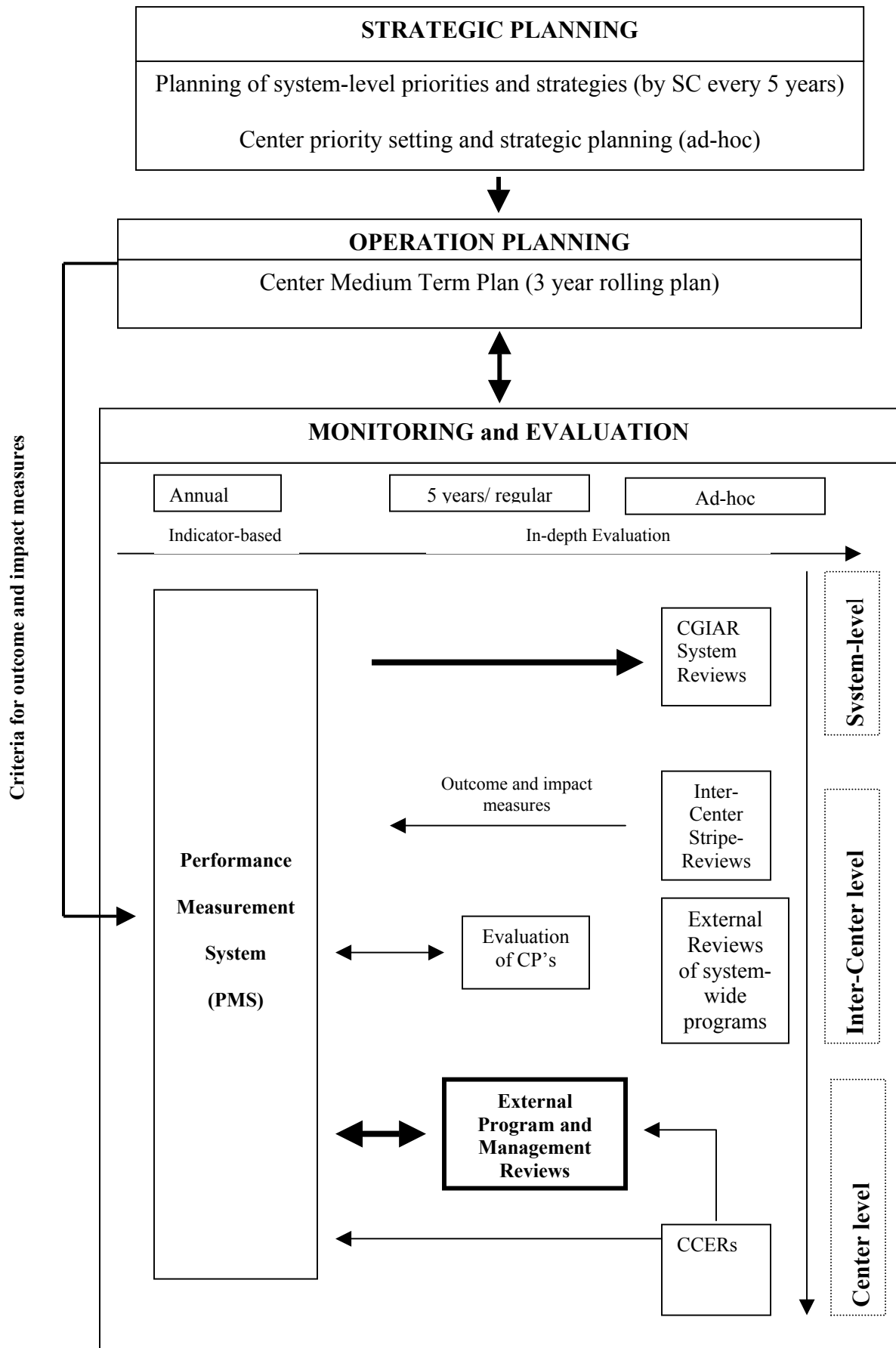
- **Inter-center thematic (stripe) reviews** - commissioned by the Science Council to evaluate specific high priority themes, e.g., Inter-Center Roots and Tubers Review; Review of Plant Breeding Methodology;
- **External Reviews of System-wide Programs** – SC-commissioned reviews of the established System-wide Programs (SWPs), e.g., System-wide Genetic Resources or Integrated Pest Management Programs;
- **Project reviews** - commissioned and/or conducted by donors at specific intervals in the project life-cycle (typically at the end of the project) or on a periodic basis, e.g., annually (e.g., donors' special project reviews or reports);
- **Center managed reviews** - i.e., internal mechanisms to ensure science and management quality control (audits, internal project evaluations, annual work-planning meetings, staff performance assessments, etc.).

The performance measurement system complements these processes by providing quantitative information to be used together with the qualitative results and by providing information annually. If centers are asked to identify their most important outputs, outcomes, and impacts annually, donors will have more up to date information in these key categories than the current five-year EPMP cycle provides, while still having the more in-depth examination and validity check of the external reviews done on longer time scales.

The Working Group expects that:

- Centers would use the results of these review processes in their performance reports, along with self-reporting, especially in the areas of quality and relevance, outputs, outcomes, and impacts.
- EPMPs and CCERs would examine the performance indicators for the center for the period of time they are reviewing as one input to their evaluation process.
- Reviewers in several of these processes may decide to use results of the PM stakeholder survey instead of conducting their own.
- Data on partnerships, outputs, outcomes, and impacts will be gathered (but not reported) at project level, so that they can also be re-aggregated easily for stripe reviews or evaluations of Challenge programs.
- Time series data from the PM System would be a valuable source of information for any CGIAR System review.

**Figure 1: Performance Measurement System in the Context of Integrated Planning in the CGIAR**



### **3. Plan for implementation**

#### **3.1 *Indicator Development and Choice***

The complexities of developing a performance measurement system and the need for consultation and buy-in at all stages, suggest that a long process will be needed. However, the Working Group discussion revealed pressures from various quarters for a more rapid development process that could run counter to these needs. These two competing pressures will have to be carefully balanced. One way of reaching this balance is to establish several “expert teams” that would further develop different elements of the system in parallel (e.g., finance, science quality and relevance, HR, etc).

The technical subgroup of the Working Group has proposed a framework for the work of these expert teams, and drawn up terms of reference for their work (Annex 2). Each team will be asked to choose three to five indicators (possibly including composite indicators) that in their judgment will fulfill the purposes of the performance measurement system best. For each indicator proposed, the team should indicate the source of the data, how it will be presented in a performance report, and a plan for establishing its reliability and validity, as well as estimating or commenting on any new costs involved in collecting it. The technical subgroup will examine the recommendations from the teams for consistency with the overall framework and with an eye to consolidating information gathering activities, minimizing costs, and making the overall set of indicators balanced and easy to understand. The team reports will be consolidated into an overall set of draft recommendations for the consideration of the Working Group. (Instructions to the specialist teams appear in Annex 2.)

The teams should be formed as soon as possible by the following lead units, since their reports are due back to the Technical Experts and Resource Persons Group by August 31, 2004.

#### **Potential to Perform**

- Quality and Relevance of Staff (SC + CDC + System Office)
- Quality and Relevance of Programs (SC + CDC)
- Governance/ Institutional Health (CBC + CG Secretariat)
- Financial Health (Center Finance Officers + CG Secretariat)

#### **Results**

- Outputs (SC + CDC)
- Outcomes (SC + CDC)
- Impacts (SC + CDC)
- Stakeholder Perceptions (Marketing Group + CG Secretariat)

#### **3.2 *Communication with Centers through Focal Points***

After this plan is sent to the CGIAR Executive Council, the System Office will ask each center by July 1 to designate a focal point responsible for providing the performance data

for the PM system. The System Office will begin communicating the upcoming expectations with the focal points immediately, by sending a short monthly update on the project, including the web site for full documentation. Since these focal points will know the most about existing data sources and competing demands for reporting within the centers, they should have a particularly strong influence on decisions about the design of data gathering systems and timing of requests. The System Office will work with CDC/CDDC to achieve good advance communication about the goals and objectives of the project, as well as to justify any additional work it may require from scientists in the system.

### **3.3 Data collection and reporting procedures**

Once the Working Group has decided on its recommendations for indicators, the specific tools needed to gather data can be constructed. A glossary or manual will be needed, with definitions of the data elements. A form for centers to use in reporting the performance data needs to be developed, as well as a format for reporting the indicators to the CG. An online reporting system is likely to be the most efficient. Regular contacts with the focal points will be needed to assure reasonable consistency in application of the definitions across centers. Standard reliability tests should be used.

*2003 as test year.* The Working Group recommends that centers be asked to enter their 2003 data into this system as soon as the system becomes available. This serves two purposes. First, it tests the procedures and allows problems to emerge early, on test data. And second, it should allow the establishment of baselines for many of the quantitative performance indicators.

*2004 as pilot year.* The data for 2004 should be entered into the system in early 2005. These data should be nearly complete, although it is inevitable that there will still be some holes where new data collection procedures are recommended. The Working Group thus considers 2004 data to be “pilot” data. The data should be made public, but with full caveats. All procedures for establishing reliability and validity should be tested on the 2004 pilot data, and revised as needed for application to the 2005 data.

*2005 as first year of full implementation.* The data for 2005, which will be gathered in early 2006, will be the first full data in the system, from the viewpoint of the Working Group. Every effort should be made to have any additional data collection procedures, such as the stakeholder survey, operational during 2005, so that the full set of indicators can appear in the 2006 report.

*Reporting on the data.* The system secretariat will develop a simple, short format for sharing the performance data with donors and centers. Wherever it is possible and appropriate, baseline data and trends should be shown, and performance across centers should be compared.

*Continued adjustment of the PM system.* The Working Group expects that revisions will be necessary and desirable in the system over time. We recommend that a group consisting of staff from the CGIAR Secretariat and Science Council Secretariat ask for

feedback and review the system annually. They should report on that review to the Executive Council, who may appoint a more broadly-constituted group from time to time to recommend changes.

### **3.4 Timeline**

April, 2004	Working Group meets, establishes work plan
May-Aug, 2004	Specialist teams develop/identify indicators, report back to WG technical subgroup
Sep 1-10, 2004	WG technical subgroup consolidates team recommendations into an overall System recommendation
September 12, 2004	Working Group approves or modifies recommendations of technical subgroup, reports to ExCo September 13-14.
Sep 15-Oct 31	Recommendations available for comment. Active consultation with Committee of Board Chairs and Center Directors Committee. Revisions as needed by the Working Group.
Nov – Jan 05	Design and implementation of online data collection form, with input from implementation group
Feb 2005	On-line data collection system is available; 2003 data entered immediately to test the system.
April 2005	Final data due into the system

## **Annex 1: Composition of the CGIAR Working Group on Performance Measurement (WGPM)**

### Members:

Kevin Cleaver, Chair, ExCo/FC (Co-chair)  
Luis Arango Nieto, ExCo/PC (Co-chair)  
Per Pinstrup-Andersen, Science Council  
Adel El-Beltagy, CDC  
Mortimer Neufville, CBC  
Stephan Krall, EIARD  
Selcuk Ozgediz, System Office designate  
Experts on performance measurement  
-- Flavio Avila, EMBRAPA  
-- Susan Cozzens, Georgia Institute of Technology

### Resource Persons:

Ruben Echeverria (IDB)  
Doug Horton (CIAT)  
Sirkka Immonen (SC Secretariat)  
Maria Iskandarani (CGIAR Secretariat)

## **Annex 2: Instructions to Specialist Teams in Choosing and Developing Indicators**

The Working Group has established eight performance elements for the new system, grouped into two broad performance clusters. It has also adopted dimensions of the element and definitions of key terms.

Your job is to identify three to five key performance indicators to be reported under the element you are working on. Qualitative formats for reporting are expected for Quality and Relevance of Programs, Outcomes, and Impacts, although sometimes quantitative assessments in these areas may be available to report. All other indicators should be quantitative, allowing for baseline measurement, center-specific targets for improvement, and comparison of performance across centers. Please reflect the Working Group's definitions and dimensions in your work.

The criteria you should use in choosing indicators are:

1. **Significance.** The indicator should reflect an important aspect of performance, not a trivial or peripheral one.
2. **Balance.** The set of indicators you choose for the concept you are working on should reflect the several dimensions of the performance concept.
3. **Transparency and simplicity.** The indicators should be easy for non-technical users to understand, without an undue amount of study or reading of fine print.
4. **Use of existing data.** The indicators should not call for the collection of new data or the construction of new analysis unless there is no other way to report on the performance concept (as in the case of stakeholder perceptions). The costs of collating the information should be kept to a minimum.
5. **Reliability and validity.** The reliability and validity of the indicators should be testable in normal auditing procedures. An exception here is any indicator based on expert judgment; these will be checked in existing evaluation processes.

For each indicator proposed, your team should indicate the source of the data, how it will be presented in a performance report, and a plan for establishing its reliability and validity. You should also try to estimate any new costs that might be involved.

Because of the diverse nature of Center missions and mandates, variations in their size, and the different circumstances under which each operates, what counts as good performance on an indicator may vary from center to center. Indicators should be chosen to apply to as broad a range of centers as possible. For some indicators, a minimum acceptable level of performance may be defined. For many, centers will need to set their own goals for improvement based on consultation with the System Office, the Science Council, and the center's own board.

The Working Group will incorporate your recommendations to the extent possible into a first set of performance measures to be gathered in early 2005. However, the Working Group has the ultimate responsibility for choosing an overall set of indicators that meet the criteria above, and for that reason may not be able to fully incorporate all suggestions of the specialist teams.

## **Annex 3 Description of Elements**

### **ELEMENT 1: Outputs**

#### **Concept/ rationale**

At the core of assessing “how well a center is doing” is information on the extent to which the center provides outputs that lead to outcomes and impacts consistent with its mission and objectives. “Outputs” are defined in the Logframe as the “products (tangible/ intangible) delivered by projects.” The Logframe defines five categories of output: germplasm improvement, germplasm collection, sustainable production, policy, and enhancing NARS.<sup>5</sup> Each center medium term plan defines milestones, some of which may be outputs. The types of outputs that are relevant for a particular center’s work thus depend on its mission and programs.

#### **Dimensions**

*Quantitative/ qualitative.* While the Logframe definition focuses on the content of a result (a qualitative measure), the literature on research evaluation uses the term “output” to refer to the immediate, concrete products of research activities, typically, research publications and students trained (treated quantitatively). (A list used by EMBRAPA is attached.) The PM system should allow both kinds of outputs to be reported, and should separate outputs as defined by the Logframe and connected to Logframe milestones from other outputs.

*Quantity/ quality.* Both quantity and quality of outputs are important, although quantity is easiest to include in a performance measurement system. Many of the usual science output measures, such as the number of publications in journals, can be misleading if the quality and relevance of the published research is not taken into consideration. Quality measures may be available for some output indicators.

#### **Definitions**

In order to be able to count outputs on a reliable basis, a glossary will need to be developed, and use of the categories will need to be tested for reliability.

#### **Illustrations of possible indicators**

- Scientific publications, such as articles, papers in refereed proceedings, and books. Quality measures could include impact factor for the journal of publication.
- Technical publications, such as circulars and bulletins.
- Technologies, methods, and policy techniques
- Capacity building and public awareness events

#### **Caveats**

In some other performance measurement systems, counting outputs has led to unproductive efforts to increase quantity at the expense of quality; some of those systems have abandoned counts. Not all types of outputs are of equal importance, and even within a category, outputs vary in quality and impact.

---

<sup>5</sup> CGIAR Resource Allocation, December 2000.

**Table 1 - Performance indicators and output goals of Embrapa**

Performance indicators	2000/2001 Period		2002 Goals			Proposed Goals 2003 (*)
	2000	2001	Planned	Done	Plan/done	
<b>1. Technical-scientific production</b>						
Articles in refereed journals	1228	1135	1200	1191	99.2	1114
Articles in Congress proceedings	1199	1553	1563	2068	132.3	1523
Summaries in Congress proceedings	3252	3057	2640	3729	141.2	2451
Chapters in scientific books	859	656	596	732	122.8	625
Thesis Orientation (MS and PhD)	188	204	200	244	122.0	177
<b>2. Production of technical publications</b>						
Technical circular	183	197	205	205	100.0	239
R&D bulletin	145	156	232	255	10.9.9	185
Technical recommendations/communiqué	680	712	612	584	95.4	584
Books editing	134	129	145	159	109.7	131
Periodical or Document series	364	406	402	469	116.7	390
Media Technical Articles	1088	1614	1887	2193	116.2	2005
Production systems	-	-	83	79	95.2	51
<b>3. Development of technologies, products and processes</b>						
New Varieties released	73	56	53	57	107.6	46
Varieties tested and recommended	39	44	61	100	163.9	88
Practice/Agricultural process	368	370	315	395	125.4	299
Agricultural inputs	54	59	36	73	203.8	49
Agro industrial processes	50	45	52	61	117.3	38
Scientific methodology	132	189	140	171	122.1	129
Machinery/equipment	15	15	17	17	100	12
Software	27	48	58	59	101.7	50
Strip/species	11	18	11	13	118.2	13
Monitor./Zoning/Mappings	411	417	346	424	122.5	229
<b>4. Technology transfer and image promotion</b>						
Field days	831	1026	942	1114	118.3	911
Events organization	1029	1187	1022	1355	132.6	1006
High school trainees (hours)	530,060	674,363	537,956	608,529	113.1	490,199
Undergraduate trainees (hours)	1,234,450	1,463,670	1,272,945	1,626,270	127.8	1,280,069
Graduate trainees (hours)	563,231	625,555	554,823	648,521	116.9	476,608
Courses offered (hours)	21,392	27,079	20,847	27,525	132.0	20,842
Folders Printed	328	370	427	520	121.8	457
Videotape production	300	398	475	727	153.1	429
Demonstration and observation units	5,712	3,853	3751	4,174	111.3	3,462
Seminar presentations	7,633	8,965	7,914	11,696	147.8	7,838
Technological press reports	4,836	7,169	7,092	10,197	143.8	8,111

(\*) - Goals not yet negotiated with the Embrapa board.

## **ELEMENT 2: Outcomes**

### **Concept/ rationale**

Before the results of CGIAR research can achieve the system's goals and produce benefits for the public, they must be incorporated into the practices of someone outside the centers themselves. The concept of "outcomes" refers to those steps that someone outside the system takes to translate CGIAR results into actions that lead toward benefits. Each Center's medium term plan defines milestones, some of which may be outcomes.

### **Dimensions**

Outcomes can be categorized several ways:

- in relation to the CGIAR's three system goals, poverty alleviation, food security, and sustainable food production
- in relation to the CGIAR's five Logframe output categories: germplasm improvement, germplasm collection, sustainable production, policy, and enhancing NARS.<sup>6</sup>
- in relation to CGIAR's five budget categories: increasing productivity, strengthening NARS, protecting the environment, improving policies, and saving biodiversity
- in relation to the goals of individual centers, set in response to five-year reviews.

Centers could be asked to report outcomes in any of these categories. As long as the choice of categories is consistent across the performance measurement system, performance reporting would allow aggregation of outcomes at the system level.

### **Definitions**

In the terms of the Logframe, outcomes are "purposes," "utilization of the CGIAR outputs by those who receive them."

### **Illustrations of possible indicators**

Typically the extent of use of scientific research has been judged qualitatively through expert opinion provided by peer reviews such as EPMRs and CCERs. There has also been some development of quantitative measures of short term outcomes, such as the volume of research stimulated elsewhere. These should be reported when they are available. Examples of outcome indicators from the Logframe include:

- CGIAR technologies and methods utilized within NARS programs
- Options and consequences from CGIAR policy and public management research are taken into consideration by decision-makers
- Options employing CGIAR outputs have been adopted and codified in the form of rules, regulations, laws, etc.

---

<sup>6</sup> CGIAR Resource Allocation, December 2000.

Other research organizations frequently report on use of their results through examples, and the CGIAR performance measurement system expects to receive qualitative reports on outcomes, with an annual summary of not more than one page. See also the discussion of measures for impacts.

### **Caveats**

There is often a considerable time lag between the completion of scientific research and the development of visible outcomes, making annual measurement and reporting difficult. Some techniques have been developed to mitigate these challenges. Periodic surveys of clients have been used to obtain estimates of outcomes. Some organizations have reported annually on the average of outcomes achieved over a several year interval (“moving averages”). For example, the U.S. General Accounting Office tracks the response to its recommendations over several years and reports annually on accomplishments of work completed in the previous three years.

## **ELEMENT 3: Impacts**

### **Concept/ rationale**

The goals of the CG system -- alleviating poverty, increasing food security, and improving sustainability -- express the impact it wishes to have in the world. Impacts are thus the most important indicators of performance of the individual centers and of the system as a whole. If centers cannot demonstrate their impacts, good performance in the other elements is meaningless.

### **Dimensions**

See discussion under Element 6, Outcomes.

### **Definitions**

“Impacts” refer to the longer range social, environmental and economic benefits that are consistent with the center’s mission and objectives, e.g., increased agricultural productivity, improved food distribution, etc. Under the Logframe, centers are accountable for achieving “intermediate goals,” that is, direct benefits for target groups resulting from the uptake of innovations (outputs). The intermediate goal of the system is phrased as:

Productivity of resources in Agriculture, Fisheries, and Forestry is increased and the sustainable management of natural resources is improved.

Centers should be able to provide a narrative summary of their “impact pathways,” by specifying beneficiaries and types of benefit with consequences for near and future productivity as well as the sustainable management of natural resources. They should also be able to summarize center contributions to achievement of CGIAR goals.<sup>7</sup>

### **Illustrations of possible indicators**

The impacts of science are particularly difficult to assess. They are often influenced by a number of factors, with science playing a limited contribution. As a result, it is difficult to attribute changes in these areas to any particular scientific input. Time lags for their realization can be even longer than for outcomes, perhaps taking more than a decade. The assessment of impacts often requires complex evaluation studies, which can only be conducted from time to time. The impacts of science can also differ widely in terms of their importance. Some breakthroughs may be so significant that in and of themselves they justify the resources used by a center. It is important that the performance measurement system reflect these differences and that centers be given the opportunity to highlight these “significant accomplishments”.

The problem of attributing results to a particular action is often a problem for measuring outcomes as well as for measuring impacts. In the absence of clear evidence for attributing results, it is important to establish a clear logical link. A basic principle of

---

<sup>7</sup> Manual on Logframes within the CGIAR System, December 1999.

performance measurement is that it should provide a picture of how outputs lead to outcomes which in turn contribute to impacts consistent with the center's mission and objectives. A common practice is to develop a framework that depicts these relationships and provides the basic direction for selecting performance indicators for results. The current CGIAR Logframe could serve as a starting point for developing indicators and demonstrating linkages among outputs, outcomes and impacts.

The work of the SC Standing Panel on Impact Assessment is an important source for the development of indicators of outcomes and impacts. Developing indicators of outcomes and impacts will in all likelihood involve finding the proper balance between qualitative and quantitative assessments of impacts. Even when qualitative judgments are reached it is important they have a sound factual basis.

Indicators will also have to take into consideration the differing mandates of individual centers. Not all centers will be able to show achievements with regard to germplasm development and conservation, for example. Others may not conduct policy research. Capacity building of NARS, for example, is an objective of all centers. Indicators will have to be developed in such a way that a common level of assessment is available while allowing sufficient flexibility so that the assessment of a given center's accomplishments reflect its mission and mandate.

**Illustrative Indicators:**

- Economic
- Social
- Environmental
- Institutional

As one moves down the list, less research has been done and fewer measures are available. Possible information sources include center records, external CGIAR reviews, CCERs, and impact assessment studies.

**Caveats**

Impacts cannot be produced on an annual schedule, and many of them cannot be quantified. Reporting will need to be qualitative, and may need to take place less regularly than the annual schedule for the other elements. The Working Group envisions a summary of impacts for each center of a page or less.

## **ELEMENT 4: Stakeholder Perceptions**

### **Concept/ rationale**

The external image of the centers and the CGIAR among donors, partners, and other stakeholders is very important for continued public support for their objectives and activities. This support contributes to the establishment of external partnerships and to the take-up of CGIAR and center research products, thereby contributing to the outcomes and impacts to be derived from these products.

### **Dimensions**

Stakeholders have impressions of all aspects of center performance, from quality through management to impacts.

### **Definitions**

Stakeholders include donors and partners for CGIAR centers, along with users of center results and public organizations that show an interest in center work.

### **Illustrations of possible indicators**

To include stakeholder perceptions in the performance measures on a systematic basis, a new survey will probably need to be developed and administered. This would not need to be done on an annual basis. Content analysis of published statements about the nature of the centers/CGIAR could also be tested as a useful source of perceptions.

### **Caveats**

Stakeholder perceptions are an important source of measurement of the influence of what the centers do. But because stakeholder perceptions are not always informed by facts, there is a risk that the survey results will sometimes carry more influence than they should. A new survey would be costly, and those costs should be carefully weighed against the benefits of this performance element.

## **ELEMENT 5: Quality and Relevance of Staff**

### **Concept/rationale**

The CGIAR Centres engage in scientific research and research related activities. They maintain an image of “Centres of excellence” which is essential for securing a credible position in the international agricultural research continuum where the Centres’ niche is primarily in conducting strategic and applied research. Maintaining high scientific quality in staffing is a prerequisite for maintaining visibility in the scientific community and credibility among donors and stakeholders.

The CGIAR Centres are highly mission oriented and each through its own mandate and goals contributes to the CGIAR’s overarching goal to reduce poverty and protect natural resources in order to achieve sustainable food security. Therefore the match between staff skills and centre mission is an equally important prerequisite for effectiveness.

### **Dimensions**

This element has at least two dimensions: the quality and relevance of staff at any given time and the measures in place to recruit and retain staff capable of high quality, mission-oriented research.

### **Definitions**

High quality implies that research is conducted following best scientific principles which in general include self-criticism, openness and communication and specifically include use and development of the most appropriate methodologies of investigation and validation in each field of research. Quality also implies innovativeness.

Relevance implies that the CGIAR’s general criteria for prioritization are fulfilled, i.e., comparative advantage viz. a viz. alternative sources of supply, probabilities of success and cost-effectiveness, international public goods nature of the results and contribution to the CGIAR’s overarching goals. Those goals are in turn chosen in relation to at least two factors: the beneficiaries’ and partners’ needs (demand); and the opportunities provided by advances in research (supply). CGIAR research should be relevant to some agro-ecological or geopolitical region beyond the individual site or country, i.e., provide regional or global public goods.

### **Illustrations of possible indicators**

Quality and relevance are best judged through expert review processes involving scientists and users, using a base of objective information about programs and accomplishments. In the CGIAR system, the five-year EPMRs are the main process that makes these judgments. CCERs are sometimes used to supplement them. Results of the most recent reviews of this sort can be reported in the performance measurement system on a qualitative basis. Over time, the Science Council may wish to design a standardized rating scale for external reviewers to use in summarizing their results to go into the performance measurement system.

In the meantime, however, proxy indicators that can be gathered on an annual basis may be useful in this area. For annual performance monitoring purposes, it may be possible to adopt short term, limited quantitative indicators of staff quality such as the number of honors and awards and other concrete signs of recognition from professional communities. Such indicators, however, only supplement and do not supplant expert judgments.

## Quality

Proxy indicators for quality can be sought from following areas of parameters that reflect recognition, visibility and linkages. Especially useful are those that imply peer assessment of the quality dimension.

Parameter	Source/Indicator
<b>Individual indicators</b>	
Invitation to keynote lecturing	Reports from scientists
(Significant) honors, prizes and awards	Reports from scientists
Membership in editorial boards	Reports from scientists
Reviewing for journals	Reports from scientists
Membership in panels/boards	Reports from scientists
Stewardship of academic studies (MSc, PhD, Postdoctoral fellows; lecturing)	Reports from scientists
Diversity of staffing (gender, ethnicity, age, region)	
<b>Processes to maintain staff quality</b>	
Success in recruitment	quality of candidates, getting first choice candidate
Staff assessment, and recognition of scientific excellence	
Team assessment	
Composition of research teams (disciplinary, age, seniority)	
Evaluation process that facilitates pre and post evaluation consultation	
% budget spent on HR development	

## Caveats

Quality of individual may not reflect optimal team functioning ability. Individual scientific esteem may be contradictory to what is required for effective interdisciplinary team performance.

## **ELEMENT 6: Quality and Relevance of Programs**

The CGIAR Centres engage in scientific research and research related activities. They maintain an image of “Centres of excellence” which is essential for securing a credible position in the international agricultural research continuum where the Centres’ niche is primarily in conducting strategic and applied research. Maintaining high scientific quality in research is a prerequisite for maintaining visibility in the scientific community and credibility among donors and stakeholders. It is a prerequisite for attracting relevant partners and for competing for funds increasingly distributed through competitive schemes.

The CGIAR Centres are highly mission oriented and each through its own mandate and goals contributes to the CGIAR’s overarching goal to reduce poverty and protect natural resources in order to achieve sustainable food security. Therefore the relevance of research and related activities is an equal to quality as a prerequisite reaching the goals and fulfilling the mission effectively.

Both scientific quality and relevance of programs provide an indication of the extent to which Centres are capable of conducting productive research toward achieving the agreed goals and fulfilling the mission.

### **Dimensions**

The element has at least two dimensions: the quality and relevance of on-going research, and the measures in place for securing quality and relevance of future activities. Strategic planning and priority setting processes are the main tool for securing relevance. Feedback through reiterative planning, including *ex ante* impact assessment, and stakeholder/partner consultation inform the centre of the relevance of on-going activities.

### **Definitions**

High quality implies that research is conducted following best scientific principles which in general include self-criticism, openness and communication and specifically include use and development of the most appropriate methodologies of investigation and validation in each field of research. Quality implies also innovativeness.

Relevance implies that the CGIAR’s general criteria for prioritization are fulfilled, i.e., comparative advantage viz a viz alternative sources of supply, probabilities of success and cost-effectiveness, international public goods nature of the results and contribution to the CGIAR’s overarching goals. Those goals are in turn chosen in relation to at least two factors: the beneficiaries’ and partners’ needs (demand); and the opportunities provided by advances in research (supply). CGIAR research should be relevant to some agro-ecological or geopolitical region beyond individual site/country, i.e., provide regional or global public goods.

### **Illustrations of possible indicators**

Quality and relevance are best judged through expert review processes involving scientists and users, using a base of objective information about programs and their

accomplishments. In the CGIAR system, the five-year EPMRs are the main process that makes these judgments. CCERs are sometimes used to supplement them. Results of the most recent reviews of this sort can be reported in the performance measurement system on a qualitative basis. Over time, the Science Council may wish to design a standardized rating scale for external reviewers to use in summarizing their results to go into the performance measurement system.

In the meantime, however, proxy indicators that can be gathered on an annual basis may be useful in this area. For annual performance monitoring purposes, it may be possible to adopt short term, limited quantitative indicators of science quality and relevance at team or center level, such as success in competitive grants schemes. Such indicators, however, only supplement and do not supplant expert judgments.

Parameter	Source/Indicator
<b>Indicators of ongoing quality</b>	
Success in competitive grants schemes	Center records
<b>Processes to encourage quality</b>	
Internal programme/ project reviews, e.g. annual programme meeting	Available/quality of process/conclusions
<i>Ex ante</i> project reviews (internal, external)	
External peer reviews of research	CCERs
Innovation funds to promote high risk/potential high impact research or for piloting/scoping	
Externally commissioned external reviews	
Mechanisms for rewarding teamwork	
Regular technical seminars held	
<b>Indicators of ongoing relevance</b>	
Stakeholder funding share in center activities	Financial records
Quality partnerships toward MTP milestones	Researcher reports
<b>Processes to encourage relevance</b>	
Priority setting	
Stakeholder consultation and involvement	
Scientific opportunity mapping	

## Caveats

While there are several relatively unambiguous indicators of quality, peer assessment is often needed for more accurate assessment of quality. Creativity and innovativeness are not planned or programmed. For the measures of partnership and collaboration, an increasing level of partnerships is probably more important than the overall level, so this may not be a good measure to compare across centers. Small centers have fewer opportunities to partner than large centers, so any measure would need to be adjusted by size. Quality and productivity of partnerships is more important than quantity, but no convenient measures are available. Partnership measures in some other systems have led to the proliferation of empty partnerships, an activity that drains energy from more substantive work.

## **ELEMENT 7: Governance and Institutional Health**

### **Concept/Rationale**

The CGIAR, its affiliated centers, and those it is intended to serve – principally the rural poor – operate in complex, rapidly changing environments. For this reason, the goals, strategies, activities and relationships of research centers and other actors in innovation systems are changing at an accelerating pace. In this context, continuous learning and change are needed for the CGIAR to survive and for it to continue to make meaningful contributions to the broad goal of sustainable poverty alleviation.

Continuing success and improvement depend upon a center being effectively governed and on the health of its internal institutional climate. In a performance measurement system it is important to have indicators that will demonstrate the extent to which the organization's governance system fosters performance and institutional health. It is important to document and monitor the extent to which the organization has in place the key components and processes that will contribute to and sustain institutional health and viability.

### **Dimensions**

Governance. The health and viability of the center will be affected by the extent to which those who govern are knowledgeable of the centers' mission, goals and priorities. Appropriate structures must be in place to ensure effective stewardship of the center's resources. Measures should be taken to ensure the involvement of staff in program planning and management decisions.

Institutional Health. There must be clear understanding of the center's long and medium-term plan and thus consistency with the centers' mandate, mission and goals. Measures must be taken to ensure appropriate staff job satisfaction. There should be assurances that grievances can be filed and resolved without reprisals. Program evaluation should be used for program improvement.

Culture of Learning and Change. The center's culture should promote collaboration and risk taking, and reward teamwork and the achievement of organizational goals. The center should have effective internal and external communication and knowledge management processes. Staff should be entrusted with responsibilities that will ensure capacity building and opportunities that will enhance their leadership skills.

Partnerships. Partnerships with other centers and external organizations allow centers to accomplish more than they could by themselves by complementing existing human and financial resources. Increased involvement with other CGIAR centers enhances progress toward common goals, while involvement of external organizations extends the reach of center activities.

### **Illustrative indicators – Governance**

- System in place for recruitment and retention of top managers
- Existence of managers with a proven record of financial, personnel and management skills.

- Appropriate size and composition for the Boards of Trustees, and relevant sub-committee structure.
- Board members' orientation program and a well defined total board and individual board member self evaluation system.
- Effective performance of audit and program oversight functions.
- Demonstrated knowledge of CGIAR donor community
- Number of organizations and individuals involved in periodic planning and review exercises.

### **Illustrative indicators – Institutional Health and Culture of Learning and Change**

- Existence of job satisfaction, self improvement and leadership development plan
- Frequency of changes in organization structure
- Existence of an appropriate reward system and a clearly defined grievance process
- Conduct of EPMR and CCER in a timely manner with appropriate relevant and accurate responses to evaluators' recommendations and concerns
- Center and Board commitment to diversity
- Mechanisms for rewarding innovation and teamwork.

### **Illustrative indicators – Partnerships**

- Stakeholder funding share in center activities.
- Inter-institutional collaboration in the conduct of research projects, production of technical documents, or development of technologies products and processes

Information Sources: MTP, Center Annual Reports, Board of Trustees Evaluations, Employee Surveys, Staff Evaluation Plan, Grievance Procedure document, External Reviews and Reports, Audit Reports, Board Orientation Plan, Recruitment Plan

The Technical Group discussed the possibility of using a checklist format to gather information on these processes. All these should be in place. There should be no new costs except for compilation of information.

### **Caveats**

Quality of the processes is not easily incorporated into a performance measurement system. There is a risk that diagnostic surveys without adequate follow-up action could lead to frustration and reduced staff morale.

## ELEMENT 8: Financial Health

### Concept/Rationale

Financial health – the effective acquisition and management of financial resources - is fundamental to a center’s performance. For an organization to make reasonable progress towards its goals and objectives, it must have the required resources available when needed. Sufficient flexibility in its financial situation helps an organization be innovative, take advantage of new opportunities and adapt to changing circumstances. Reporting on financial health is also a fundamental aspect of any accountability relationship and helps funding organizations to make resource allocation decisions.

### Dimensions

- Short term solvency (liquidity)
- Long-term financial stability (adequacy of reserves)
- Operating leverage (management of fixed costs)

### Definitions

- *Short term solvency (liquidity)*. The cushion of liquid funds available to the Center to: (a) continue normal operations during the year despite the “lumpiness” experienced in actual receipt of donor funds; and (b) provide sufficient time to Center management to initiate special cost saving, income generating or short term financing measures in response to unexpected events affecting the cash flow of the Center. In short, debt-paying capacity of the center.
- *Long-term financial stability (adequacy of reserves)*. The Center’s financial capacity to smoothly manage changes to activity levels over the longer term in the event of shocks affecting income or expenditures.
- *Operating leverage (management of fixed costs)*. The degree of flexibility that the Center has to adjust to changing financial conditions.

### Illustrative indicators

- *Short term solvency (liquidity)*: **Number of days working capital**. (The Center Directors recommend 90 to 120 days as an acceptable range.)
- *Long-term financial stability (adequacy of reserves)*: **Number of days unrestricted net assets adjusted for fixed assets**. (The Center Directors recommend 75-90 days as an acceptable range.)
- *Operating leverage (management of fixed costs)*: **Variable costs / total expenditures** or **Variable costs/fixed costs**.

⇒ *Sources*: Annual center financial statements. Some definitional issues will need to be worked out among the centers (e.g., what expenditures shown in the financial statements classified as variable, whether all fixed costs or a selected set should be used , etc..)

These have already been agreed by financial experts and two are endorsed for use by the center directors.