Towards Designing a Performance Measurement System for the CGIAR

Annexes

CGIAR Working Group on Performance Measurement

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)
Michael Cernea, TAC/iSC
Adel El-Beltagy, CDC
Mortimer Neufville, CBC
Stephan Krall, EIARD
Selcuk Ozgediz, System Office designate
Experts on performance measurement
-- Flavio Avila, EMBRAPA/Yale Univ
-- Stan Divorski, Consultant

Resource Persons:

Ruben Echeverria (IDB)
Doug Horton (ISNAR)
Sirkka Immonen (iSC Secretariat)
Maria Iskandarani (CGIAR Secretariat)

Annex 2

Background on Global Trends, Concepts and Lessons Learned in Performance Measurement

A. Global Trends In Performance Measurement

In the U.S., the Government Performance and Results Act of 1993 requires all federal agencies to set performance goals and to report on annual performance compared with the goals. The Act was an effort to improve government effectiveness and efficiency by providing agencies with the tools and incentives to manage their agencies for results and to improve accountability by providing Congress and the public with clear information on what government programs were accomplishing with tax dollars. What some observers initially thought to be a short term initiative has in fact turned out to be an enduring effort on the part of several successive Congresses and two presidential administrations, supported by both major political parties.

In 1995, Canada revised the administrative requirements for federal agency budget reporting to require information on performance targets and accomplishments. Currently, major Canadian agencies are required to publish annual performance expectations every spring and a report of progress against the previous year's expectations each fall. In both Canada and the U.S. each government's treasury (the Office of Management and Budget in the U.S. and the Treasury Board in Canada) provides an annual report on overall government performance against expectations.

In 1988, the United Kingdom established executive agencies to deliver government services, separating that function from policy. These agencies operate under a greater degree of flexibility than policy departments, making accountability an issue. In order to address accountability, these agencies publish Annual Performance Agreements and Annual Reports, respectively containing information on their expected and actual performance against specific targets. In the UK, performance reporting extends down to the local government level, where performance against a wide range of established performance indicators for municipal services such as police, fire control, social services, and libraries is published annually in a comparative format by the British Audit Commission.

In the Australian federal government, program management and budgeting reforms in the mid-1980s led to the inclusion of performance information in Portfolio Budget Statements, which provide parliament with budget information on policy groupings of agencies. The purpose of the performance information is to better enable Parliament to understand budget requests. Performance information is also required in the Annual Reports of individual departments and agencies. These annual reports are intended to demonstrate individual agency accountability to Parliament.

Developments are not restricted to English speaking countries. Sweden has required government agencies to report results annually since the mid 1980's. The Netherlands has recently put in place regulations that require government agencies to provide ongoing monitoring and evaluation information with respect to a number of performance issues, including the extent to which policy targets are being met and the appropriateness of the means being used to address these targets. Recent legislation in Spain requires "result monitoring and management" to improve management and the internal operations of government organizations. Neither are these developments restricted to national governments The Canadian provinces of Alberta, Nova Scotia and British Columbia have long standing performance reporting requirements, as have U.S. states such as Oregon, Florida, Washington, Texas, Missouri and North Carolina. Local governments, most notably in the U.S. have been moving in this direction as well, in certain cases in advance of federal efforts in the area.

B. Trends in Research Organizations

National Research Centers

The National Institute for Health (USA) and the Cooperative Research Centers Program (Australia), both respond to the requirements made by their respective governments.

- ***National Institute of Health (NIH) in the USA***
  The National Institutes of Health (NIH) is the federal focal point for medical research in the United States. The NIH, comprising 27 separate Institutes and Centers, is one of eight health agencies of the Public Health Service which, in turn, is part of the U.S. Department of Health and Human Services.
  Their performance measurement system is very much target oriented and is a response to the Government Performance and Results Act (GPRA). NIH's annual performance plans include both 1) performance goals that can be assessed through quantitative measures and 2) performance goals based on descriptive achievement criteria. Where quantitative measures can be used, performance assessment is a process, principally, of comparing data on actual achievement with the target levels stated by the Annual Program Performance Plans. In addition, GPRA allows an alternative approach as a way for an agency to identify performance goals based on criteria which are chiefly descriptive in nature. For the evaluation of some of NIH's output goals an independent review process has been put in place.

- ***Cooperative Research Centres Program - Australia***
  In May 1990, the Australian Commonwealth Government launched the Cooperative Research Centres Program. The Program currently supports 62 Cooperative Research Centres (CRCs), which are collaborative research ventures bringing together researchers from universities, the public sector and business. The CRC performance measurement framework, currently under approval, has a strong focus on outcome, which should help to identify: (1) the ***impacts*** that are

the manifestation of its outcomes (2) the ***effectiveness*** of the Program in achieving outcomes that address its objectives (3) the ***appropriateness*** of the CRC Program objectives to meeting national and community needs (4) the ***efficiency*** with which the program uses inputs to produce its outputs. This conceptual framework has been translated into a set of key information and performance indicators, that are measured annually.

Agricultural research institutions

There are also examples for performance measurement in the agricultural research sector:

- ***Brazilian Corporation for Agricultural Research (Embrapa) – Brazil***
  A model developed by EMBRAPA deserves careful consideration by CGIAR as it is one of the most advance models for performance measurement by a scientific institution. For this reason, a lengthier description of the model is attached as Annex 3. The Brazilian Corporation for Agricultural Research (Embrapa) created an 'Evaluation and Rewarding System based on Results (SAPRE)' to increase the productivity of the Corporation research centers, in the fulfillment of its institutional mission of generating and spreading technologies for the Brazilian agribusiness (Portugal et. al., 1996). The focus of the system is to reward (through wages add ins, advantages in the allocation and availability of funding, and by non-monetary prizes) the units, work teams and employees that most contributed to the accomplishment of the mission and the objectives of the Corporation. It should stimulate a healthy competition and cooperation among units, teams and individuals, in such a way to become them more efficient. The institutional evaluation system SAPRE comprises 7 basic components: (1) Accomplishment of institutional goals (efficacy); (2) Economic efficiency; (3) External raising of funds; (4) Quality auditing; (5) Impact evaluation; (6) Image assessment; and (7) Productivity. Development of the system began in 1995.

- ***Agricultural Research Service (ARS)***
  The ARS is the principal in-house research agency of the U.S. Department of Agriculture (USDA). ARS Research is organized into 22 National Programs (NP) on agriculture, nutrition, technology, the environment and other topics that affect the American people on a daily basis. Again in response to GPRA, the ARS is conducting annual program performance evaluations and tracks progress against approximately 250 "indicators of progress" organized under the goals and performance goals from the ARS Strategic Plan.

University based research

There are also examples from Universities, who increasingly get under accountability pressure and their funds are more and more allocated according to performance measurers. Examples for performance–based fund allocation are:

- ***Higher Education and University Research in the UK***

The so called "Research Assessment Exercise (RAE)" is a performance measurement system for higher education that has been in place in the UK since 1986. The purpose of the RAE is to provide ratings of the quality of research conducted by universities and higher education colleges in the UK, and based on this rating system, to make allocation decisions on funds.

- ***Higher Education System - Australia***
  The Australian government recently reformed the funding scheme for research in the higher education system. On pillar of the funding scheme is a performance based system for block funding of universities for their research and training. The performance indicators used are related to attracting (1) research students, (2) research income, and producing (3) the output in terms of research publication.

- ***Tertiary Education System - New Zealand***
  The government of New Zealand is in the process of establishing a Performance-Based Research Fund (PBRF). Funds will then be allocated according to an aggregated "quality score" achieved by an institute, which is calculated from an overall performance measurement of individual researchers (indicators are their total output; peer esteem factors (such as awards, fellowships), and their contribution to the development of new researchers and/or a vital high-quality research environment). Additional indicators are the number of research degree completed at the institute and the amount of external research income gained by the institute.

C. Conceptual Frameworks

As the notion of performance measurement and management emerged over the past 10 to 15 years, various conceptual models of performance management systems were developed. The sourcebook on performance measurement in research institutions and programs (CGIAR Secretariat, 2003), which was shared with the WG, describes four major conceptual models that currently find practical application in the private and public sector. These approaches all share a broad view of performance measurement that includes information not only on organizational accomplishments, but also on the capacities needed to provide deliverables that meet the needs to the intended clientele. These conceptual models will be shortly summarized in the following – for more details please refer to the sourcebook.

Common Assessment Framework (CAF)

CAF was developed in 2000 and is a result of the cooperation among the EU Ministers responsible for public administration. It has been designed for use in all parts of the public sector, and to be applicable to public organizations at all administrative levels.
CAF is designed in a simple manner, suitable for providing an initial impression of how the organization performs. Among its purposes are to serve as a tool for public service managers and to facilitate benchmarking between public sector organizations. The

framework basically consists of two groups of evaluation criteria:[1]

(1) Enablers criteria: *Leadership*
Strategy and Planning
Human Resources Management (HRM)
Partnership and resources
Process and change management

(2) Results criteria: *Customer-oriented results*
People results
Society results
Key performance results

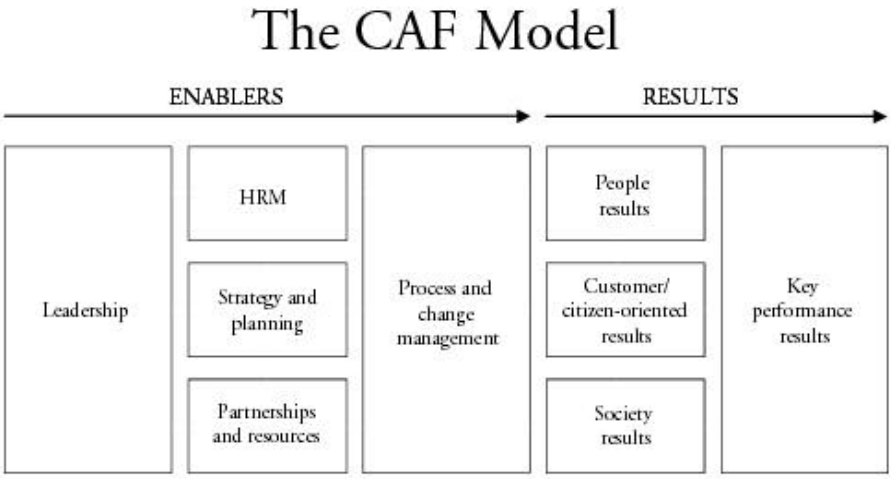

Figure 1: The CAF Model

The Balanced Scorecard

The concept of 'Balanced Scorecards' was first introduced by Kaplan and Norton (1992). The BSC is not simply a set of performance measures; rather it is an approach for a strategic management system. BSC is intended to:

- Capture the complexity of all activities and processes within an organization and to cascade responsibility for performance in a transparent way down to the individual employee
- Make visions and derived strategic goals measurable
- Better reveal strategic goals to the employees
- Link Strategies and activities to budgets
- Make it easier to adapt strategies to a changing environment.

The key feature of the balanced scorecard is that it supplements traditional financial measures with criteria measuring performance from three additional perspectives:

[1] The CAF Model is inspired by the EFQM Excellence-Model (European Foundation for Quality Management)

- **Financial Perspective** – "To succeed financially, how should we appear to our shareholders?"
- **Customer Perspective** – "To achieve our vision, how should we appear to our customers?"
- **Internal Business and Process Perspective** – "To satisfy our shareholders and customers, what business process must we excel at?"
- **Learning and Growth Perspective** – "To achieve our vision, how will we sustain our ability to change and improve?"

For each of the perspectives the institution has to define *objectives, measures, targets, milestones and activities*. Thereby the organization's high level strategic objectives and measures can be translated into objectives and measures for operating units and even individuals. The idea behind this concept is to enable companies to track financial results while simultaneously monitoring progress in building the capabilities and acquiring the intangible assets they would need for future growth (Kaplan and Norton, 1996).

Kaplan and Norton (2000) adjusted the original architecture of the BSC to the NGO and public sector environment, who typically do not view the financial perspective as a primary objective. Furthermore, in case of the private sector, the customer both pays for and receives the service. But in a nonprofit organization, donors provide the financial resources (or pay for the service), while another group receives the service. Kaplan and Norton (2000) suggest placing an overarching objective at the top of the scorecard that represents a long-term objective (see figure 2).

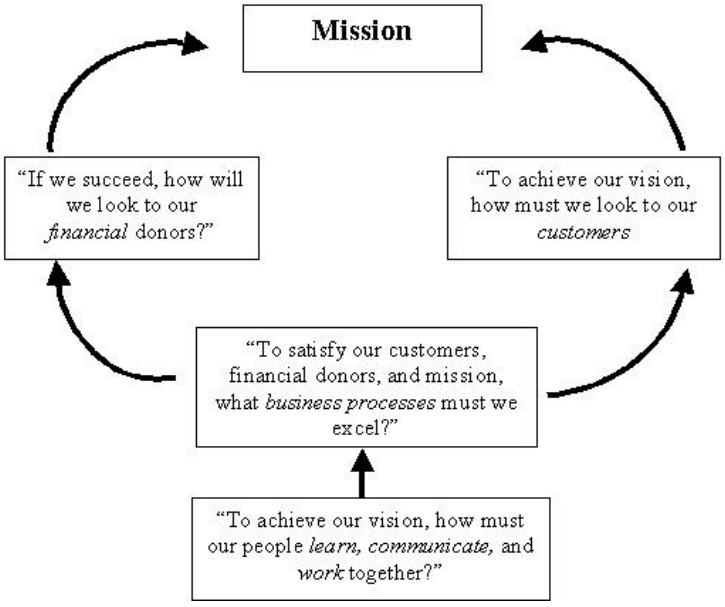

Figure 2. Adapted BSC Framework to Non-Profit Organizations (Kaplan and Norton, 2000)

Total Quality Management (TQM)

Total Quality Management is a philosophy that came out of the pioneering work of Deming and Juran in Japan in the 1980's. At first glance TQM is seen as a change in an organization's technology, i.e., its way of doing work. On the service side, this means the way clients are processed the service delivery methods applied to them and ancillary organizational processes such as procurement, order processing, etc. But TQM also focuses on change in an organization's culture: its norms, values, and belief systems about how organizations function. And finally, it focuses on change in an organization's political system: decision making processes and power bases. The conceptual idea is that for substantive change to occur, changes in all three dimensions must be aligned.

In other words, TQM is a management led process that actively involves every employee in satisfying customer needs (internal and external) by continuously improving all aspects of work activity through structured control, improvement, and planning methods. It requires a transformation of the roles of Management and of employees empowering them to continuously improve their work processes. Deming outlined 14 steps that managers in any type of organization can take to implement a total quality management program.

ISO 9000

The International Organization for Standardization (ISO) is a worldwide federation of national standards bodies. The ISO9000 family of standards has been developed to assist organizations, of all types and sizes, to implement and operate effective quality management standards. ISO 9000 specifies *requirements* for a quality management system for any organization that needs to demonstrate its ability to consistently provide product that meets customer and applicable regulatory requirements and aims to enhance customer satisfaction. The ISO9000 family of standards follow eight principles: *customer focus, leadership, involvement of people, process approach, system approach to management, continual improvement, factual approach to decision making, mutually beneficial supplier relationships.*

D. Lessons learned on Performance Measurement

A number of lessons have been drawn from experiences with PM, which can inform the design and implementation of more successful and cost-effective systems.

Start with agreement on the purposes of performance measurement. Performance information has been used for a variety of purposes, and it is important that there is discussion leading to consensus on the purposes to be pursued in the development of a PM system. It was conceive as a *mechanism to improve transparency* and *accountability for* organizational results. Governments have also implemented PM with the intention of using it as a tool to *support resource allocation* decisions and to provide incentives for improving organizational performance. PM systems have been most successful when they have been designed to serve as a *tool for management decision-making*, helping to improve programs and support strategic directions. It has also served as a *benchmarking tool*, permitting comparisons that contribute to organizational learning and the sharing of

best practices. At the same time, experience has been that trying to accomplish too many purposes will overburden the system, making it unwieldy.

Carefully manage the development and implementation of the system. Perhaps the most important single lesson is that implementation of a PM system is not purely a technical issue, but a management issue that requires careful attention to the implementation process. It is important to identify the key users of the performance information and their needs and to agree on *the goals* of the PM system. It has been found important to *focus on key goals, audiences, and indicators*, and not to expect a PM system to meet multiple goals or to overwhelm the organization with too many indicators. It is also important to *negotiate key aspects of the system design and operation with the managers and technical staff* who will have to operate the system and provide useful information. Specialized training may be needed and management systems may need to be modified to operate a PM system. There needs to be careful specification and a common understanding of the elements of the system. Responsibilities need to be clearly defined and careful co-ordination is required. *Adequate time and resources need to be invested* at different organizational levels to design and test the system, prepare guidelines, to train those who will operate the system, and to establish adequate procedures for gathering, storing and processing information.

Base the PM system on organizational goals, processes and strategies. PM in a complex organizational system, such as the CGIAR, needs to reflect the hierarchy of objectives that guide the system and its affiliated centers (performance indicators that reflect the centers' inputs, activities, outputs, intermediate outcomes and ultimate impacts) as well as the processes and strategies the centers pursue in implementing their programs (indicators of the centers' capacity or potential to perform).

Establish some guiding principles and "rules of the game." It has proven valuable to discuss basic principles and rules of conduct and agree on them with participants early on. Such principles and rules can provide direction throughout the process of developing a PM system and can aid in settling disagreements among the various groups involved in or affected by the system.

Emphasize learning and performance improvement. For a PM system to produce useful information, it must be of value to the managers and technical staff of the centers and programs who implement it. For this reason, it is important that the PM system serve as an aid to learning, management and program improvement. The information generated should be shared widely and serve as a basis for reflection on objectives, activities and results.

Integrate PM with broader management and evaluation processes. It is also important that the PM system not be operated (and viewed) as one of several stand-alone activities. It should be integrated into a broader performance management system that includes PM along with other, complementary types of evaluation, and that employs the results of monitoring and evaluation to improve program implementation and the planning of future activities.

Resource allocation decisions require a depth of understanding that cannot be provided by PM alone. Less than expected results can reflect developments beyond the control of a center's managers or technical staff. Or they may indicate the need to invest *more*, not less, resources in a center in order to bring its performance up to acceptable levels. Also,

allocation decisions that only reward better performing programs have been shown to affect the quality of performance, with managers employing a variety of strategies to convey only the most positive image of their programs. Careful attention should be given to uncovering the reasons for less than expected performance and to assisting the organization to address factors that have impeded performance.

Assess the performance of the PM system, and modify and improve it over time. Effective PM systems are not designed and implemented in one shot. They need to be perfected over time. It is important to monitor the performance of the PM system as it is implemented, to periodically evaluate the system, and to modify and improve the design and implementation procedures for PM as experience is accumulated.

Combine quantitative and qualitative indicators of scientific accomplishments. Developing quantitative indicators of the performance of science has been difficult. The traditional peer review based approaches to assessing science provide credible information on the quality of science outputs and on the benefits they have produced. However, because they are time and labor intensive, they are costly and can rarely be undertaken at frequent intervals (e.g. annually) or cover all of an organization's science programs at a single time. Furthermore, their subjective nature make comparisons across programs, institutions or time difficult. On the other hand, quantitative indicators can consume fewer resources and time, to the extent that they rely on information already collected for other purposes and are restricted to a few key indicators They also allow an easier comparison across institutions and programs. However, identifying indicators that are reliable and valid has proven to be challenging. Our review of PM systems found that the optimal strategy is to combine quantitative and qualitative indicators.

Technical Challenges in Measuring Performance

Maintaining a focus on organizational goals. A general principal of performance measurement is that activities will focus on those aspects of performance that are being measured, to the detriment of aspects that are not being measured. Consequently an organization can never rely on a single performance indicator. It is important to choose a set of indicators that represents all critical dimensions of an organization's performance. It is unlikely that the right set will be chosen initially, and continuing review and adjustment of the indicators will be necessary. In addition, evaluations are advisable from time to provide a richer understanding of an organization's performance and the extent to which it is accomplishing it's overall goals.

Understanding the causes of performance. Scores on performance indicators can change for a variety of reasons. For example, outcome measures of agricultural research can vary from time to time because changes in economic conditions affect producers' abilities to adopt new techniques, changes in weather influence productivity, or instability influences production and distribution of food. Hence, improved or less than expected performance can reflect other factors than the value of a center's research program. A logic model, such as the CGIAR Logframe, can help to understand a center's performance by carefully specifying how activities are expected to lead to specific outputs, which are expected to produce specific outcomes, which in turn contribute to the desired societal impacts. A logic model cannot successfully anticipate all outcomes and impacts, positive or negative. For example, introduction of a new plant species might affect the environment or the local culture in unexpected ways. It is important that the organization conduct regular

scans of its operating environment to help ensure that these unexpected effects are detected at an early stage. Also, evaluation studies can be conducted that analyze all the factors related to specific outcomes in order to assess the impact of the program and help ensure that unintended outcomes are not overlooked.

Balancing performance measures. Ensuring a focus on the right aspects of performance requires that indicators are well chosen and carefully balanced. For example, for service programs, indicators of service quality that measured the speed of response to client requests have resulted in less attention to the completeness and accuracy of responses. A proper balance is restored by adding indicators of completeness and accuracy. In science programs, it is important that indicators focusing on the volume of research outputs and outcomes are balanced by indicators of the significance and impact of the research, ensuring that adequate attention continues to be paid to innovative but risky research with higher potential benefit. It often takes several years to find the right balance of indicators.

Ensuring the quality of performance information. The credibility of performance information is critical to its use and has been shown to be a significant issue in many jurisdictions implementing performance measurement. The U.S. Congress has noted that most U.S. federal government agencies lack the reliable data sources and systems needed to develop performance information. Congress has also recognized that problems with performance data were deep seated and resolving them would take much time and effort. Internal and external processes and controls for assessing and monitoring information quality are needed as are practices and strategies for addressing limitations in performance information. However, detecting and correcting errors will not necessarily prevent future errors and can be prohibitively expensive. Consequently, it is important to build quality into the data collection and reporting processes. For this reason, as with the implementation of a performance measurement system in general, the quality of performance information is as much a management issue as a technical issue. It requires leadership, organizational commitment, and careful management of the development and implementation of the system.

Annex 3

Synopsis of the EMBRAPA "Evaluation and Rewarding System based on Results (SAPRE)" [2]

The Brazilian Corporation for Agricultural Research (Embrapa) developed a performance evaluation system to measure the performance of its agricultural research centers. This system, in place since 1996, involves several evaluation criteria (quantitative and qualitative) where a production model using a theoretical framework based on the analysis of production frontiers (DEA) is the main component (Avila, 2002). The set criteria used in the Embrapa evaluation system are the following:

(1) Goal accomplishment;
(2) Economic efficiency;
(3) External raising of funds;
(4) Quality auditing;
(5) Impact evaluation;
(6) Image assessment; and
(7) Productivity growth.

The evaluation of the accomplishment of goals, external fund raising, economic efficiency and productivity constitute the basis of the evaluation system, and are assessed on an annual basis. The others (quality, socioeconomic impact, and image) are complementary, and their results are progressively incorporated into the system.

(1) Goal Accomplishment

This component combines an evaluation of the accomplishment of qualitative and quantitative institutional goals. The qualitative goals include national goals established for all the centers (accomplishment of internal administrative and finance rules, establishment of partnership of internal and external, accomplishment of selected prescribed dates, reduction of maintenance costs) and other goals specific to each center (improvement of selected management process). The outputs or quantitative goals are established using a set of performance indicators, which are also used for evaluating economic efficiency (see criteria 2 and table 1).

The qualitative and quantitative goals are annually negotiated between the Directors General of the centers and the Board of Embrapa, based on the mission and objectives of each center and also their performance in former years.

Based on the accomplishment of these goals an efficacy index for each center is calculated, which also considers a weighting factor for each goal.

[2] Prepared by Flavio Avila.

(2) Economic efficiency

The economic efficiency of each research center is calculated as a result of the quotient between a set of output indicators (see table 1) over their corresponding inputs. Each output indicator is weighted according to the mission and objectives of the center, previously negotiated with the Board.

The output indicators are selected as proxies to measure the annual research production of the centers and are classified according to four categories:

a) Scientific production - by scientific production we mean the publication of articles and book chapters aimed mainly at the academic world.
b) Production of technical publications - publications produced by research centers focusing primarily on agricultural businesses and agricultural production.
c) Development of technologies, products and processes - indicators related to the effort made by a research unit to make its production available to society in form of a final product.
d) Technology transfer and image - development of actions related to the effort to make its products known to the public and to market its image.

The economic efficiency model uses as input indicators **expenditures for personnel, other expenditures and the value of annual depreciation of the capital assets**.

To analyze the economic efficiency we consider a system of dimensionless relative indices for all the indicators (inputs and outputs) of the process. These are all indices based on metric data. The idea, from the output point of view, is to define a combined measure of output as a weighted average of the relative indicators (indices) in the system. The relative indices are computed for each production variable and for each research unit within a year dividing the observed amount produced by the mean per research unit. Only research units that can potentially exercise the production activity related to the production variable in question are included in the computation of the mean. We see that, within a given year, the base of our system of production indices is defined by the set of means per unit defined by the production variables. The same procedure is used for the input data (Souza et al., 1999).

Subsequently the annual performance evaluation of the research units in terms of relative efficiency is performed based on the measurement of relative efficiency using a mathematical programming model (i.e. Data Envelopment Analysis).

(3) External Raising of Funds

The Embrapa centers are stimulated to raise funds in order to reduce their dependence on funds from the National Treasury. Therefore, one of the evaluation criteria is the amount of funds externally raised.

The system considers funds directly raised from product sales, fund raising for research projects, etc. and indirect funds, raised through the development of

technological transfer activities, production of publications, organization of events, etc. total or partially financed by partners.

Based on the amount of funds raised (direct + indirect) and the total funds received by the center from the National Treasury an index for each center is built.

(4) Quality auditing

Within the quality auditing a technical committee evaluates the research units regarding the quality of the technical-scientific program using a sample of project research proposals and their respective progress report. Quality of the proposals and results are given scores on a scale from 1 to 10. The quality evaluation of the research program as a criteria in the evaluation system is not used annually.

The quality of the Embrapa research evaluation classifies articles published by researchers in indexed journals (domestic and international) using a three rank classification of these journals adopted by the Brazilian Ministry of Education. Based on the number of articles published and the quality class of the journals an weighted index for each center is built annually.

(5) Impact evaluation

The impact evaluation measures the economic effects of the Embrapa units' research results on the agro industrial complex and on consumer's welfare, and the impact on employment rates and on health. The impact evaluation also includes an impact analysis of these technologies on the environment. All main technologies developed by each center are considered within this evaluation.

This criteria is incorporated into the system based on the quality of the impact assessment report (e.g. methodology, integration) evaluated at the Embrapa head quarter. According to a scoring system, the evaluation team gives scores for the quality of the delivered report. Furthermore, it is anticipated to also include scores for the estimated value of the benefits from a period of 5 years (strategic master plan period) for each center.

(6) Image Assessment

This area of assessment tries to evaluate the image that clients and stakeholders have about the various Embrapa research centers. A survey based on interviews with clients and stakeholders helps to gain insights into this matter. The survey is analyzed statistically and an index is generated to allow a comparison and classification of units, with regard to the degree of satisfaction of its clients.

(7) Productivity Growth

In the Embrapa evaluation system the productivity criteria relates to the annual growth of productivity of each research center, which is calculated as follows:

$$IPA_{dt} = (Output_{t} / Input_{t}) / (Output_{t-1} / Input_{t-1})$$

where:

IPA_{dt} = Annual productivity index of center "d", year "t";

$Output_{t}$ = Output index, year "t";

$Input_{t}$ = Input index, year "t";

$Output_{t-1}$ = Output index, year "t-1"; and

$Input_{t-1}$ = Input index, year "t-1";

The output and input indices are built with the same indicators used within the economic efficiency criteria, only that the indices are independent (each center) and not relative as in a DEA model. Consequently, the annual productivity of each research center is compared with its own performance in past years. The input data are deflated for the comparison.

(8) Institutional Performance Index (IPI)

The partial evaluation indices developed for each of the evaluation criteria are aggregated in a single weighted index – the Institutional Performance Index – in order to allow a classification of each unit in relation to the others. The evaluation components of efficacy, raising funds, economic efficiency and productivity are calculated annually. The remaining evaluation components are assessed every three to five years.

To assure more credibility to the evaluation system, especially concerning the centers production, the outputs and the funds raised externally (direct and indirect) are annually audited . The center have to provide evidence that the quantitative output goals were really accomplished.

Based on this performance index, the payroll share of the respective unit and according to individual performance, teams and employees may be awarded additional remuneration from a so-called Rewarding by Results Fund – FPR (component B: the Institutional Rewarding System). Furthermore, prizes are awarded for outstanding performance, in the form of non-monetary advantages and benefits to be awarded to research units, and advantages in the allocation and availability of funding (Portugal et al., 1999).

REFERENCES

AVILA, A. F. D. **Avaliação de Desempenho das Unidades da Embrapa: Metodologia para 2002**. Brasília, Embrapa. Secretaria de Administração Estratégica. Agosto 2002, 27p.

PORTUGAL, A. D., AVILA, A. F. D., CONTINI, E., SOUZA, G. S. E. Sistema de avaliação e premiação por resultados. Brasília. **Revista do Serviço Público**. 49(3):59-83, 1999.

SOUZA, G. S. E., ALVES, E., AVILA, A. F. D. Technical Efficiency of Production in Agricultural Research. Budapest. **Scientometric**s. 46(1):141-60, 1999.

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	
1. Technical-scientific production						
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
2. Production of technical publications						
Technical circular	183	197	205	205	100.0	239
R&D bulletin	145	156	232	255	10.9.9	185
Technical recommendat ions/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
3. Development of technologies, products and processes						
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
4. Technology transfer and image promotion						
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
Under graduated trainees (hours)	1,234,450	1,463,670	1,272,945	1,626,270	127.8	1,280,069
Graduated 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 until not negotiated with the Embrapa board.

Annex 4

Illustrative draft of potential performance indicators and possible information sources[3]

The potential indicators listed here are for illustrative purposes. They do not constitute an exhaustive list and may not be the appropriate indicators for the purposes of measuring the performance of the centers. However, they do constitute a starting point for further development.

[3] Prepared by Stan Divorski, based on discussions of the Technical Experts and Resource Persons Sub-Group of the CGIAR WG on Performance Measurement.

ELEMENT 1: **Center Scientific Accomplishments in Relation to Mission and Objectives.**

Description: Outputs, outcomes and impacts of center science activities

Potential Indicators [4]:

Outputs:

- Accessible germplasm that meets NARS standards
- Techniques for faster, more efficient, more reliable germplasm improvement
- Germplasm conserved in accordance with standards
- Productivity-increasing, resource conserving practices are accessible
- Research methodologies that enhance development of integrated management practices
- Information on the effectiveness of polices on production, resource management, poverty and income distribution
- Tools and techniques for improved policy analysis

Outcome:

- Germplasm, improvement practices and conservation practices utilized within NARS programs
- CGIAR outputs used by NARS and extension systems to develop production systems.
- CGIAR outputs have a visible role in orienting the work of NARS and Regional Research Organizations toward the concern of stakeholders
- 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.

Impacts:

- Adoption by producers of improved technology results in higher yields, increased production, lower prices and higher wages and employment in agriculture and associated industries
- Adoption by producers of improved technology results in lower pesticide use, lower emission levels in runoff, and maintenance of insect diversity, soil organic matter and stability of yield.
- Enhanced NARS capacity

Possible Information Sources

Existing:	New:
• Center records of outputs • External CGIAR reviews • External center reviews	• Impact Assessment studies

[4] This is a sampling of possible indicators adapted from the CGIAR Logframe.

ELEMENT 2: Efficiency	
Description: Resources required to produce a unit of output	
Potential Indicators [5]: • Efficiency of Scientific Production (academic publication) • Efficiency of production of Technical Publications (publications directed at businesses and producers) • Efficiency of development of Technologies, Products and Processes. • Efficiency of Technology Transfer • Combined index of above indicators • Year to year trends in above indicators • Ratio of fixed costs to total expenditure	
Possible Information Sources:	
Existing: • Center financial statements • Self-reporting by centers	New:

[5] Based on EMBRAPA model index for Economic Efficiency

ELEMENT 3: Stakeholder Views

Description: The extent to which centers and the CGIAR maintain a positive image that fosters continued public support for their objectives and activities.

Potential Indicators:

- Client and stakeholder awareness of CGIAR[6]
- Client and stakeholder opinions on[6]:
  - Relations with CGIAR
  - The impact of CGIAR
  - The value and relevance of CGIAR programs and activities
  - The quality of CGIAR science
- Positivity of published statements about the accomplishments and quality of centers/CGIAR.

Possible Information Sources:

Existing:	New:
• Content analysis of published statements about the nature of the centers/CGIAR	• Surveys of clients/stakeholders

[6] Adapted from CAF Society Results criterion.

ELEMENT 4: Science Quality and Relevance

Description: The quality of science that is being practiced at the centers

Potential Indicators :

- Honors and awards
- Citation Indexes log
- Ratings by independent panel
- Volume of articles weighted by quality
- Stakeholder perceptions
- Extent to which the center is employing or accessing the best scientific tools and competencies[7].
- The center's potential for further impact[7].

Possible Information Sources:

Existing:	New:
• External reviews • Center annual reports	• Independent ratings by expert panels on a biennial or triennial basis • Survey of stakeholders

[7] Qualitative assessment by independent experts

ELEMENT 5: Partnerships

Description: The extent to which centers enter into partnerships in order to complement their resources and enhance performance

Potential Indicators [8]:

- Stakeholder funding share in center activities.
- Participation in the conduct of research projects
- Provision of training for CGIAR staff
- Production of communications such as brochures and videos
- Participation in the production of technical documents
- Participation in the development of technologies products and processes
- External and internal partnerships in each of the above
- Involvement of partners in center governance

Possible Information Sources:

Existing:	New:
• Self-reporting by centers • Sciences Citation Index • Memoranda of understanding and agreements describing stakeholder financial and in kind contributions • Collaborators identified in Medium-Term plans • Co-authorships on center papers.	• Stakeholder surveys

[8] Adapted from EMBRAPA model

ELEMENT 6: Governance/Institutional Health

Description: Extent to which centers are managed so as to ensure that human and financial resources are effectively directed to the accomplishment of mission and objectives.

Potential Indicators:

- The existence of:
  - An organizational culture focused on performance.
  - Self-evaluation by the Board of Trustees.
  - Processes for follow-up of internal and external review recommendations.
  - Processes to ensure effective stewardship of resources.
  - Processes for human capital management that ensure the organization has and will continue to have the appropriate level and blend of capacities for accomplishing its mission and objectives.
  - A leadership development program
  - A grievance process
- Staff job satisfaction
- Adequacy of internal audit
- Re-occurrence of CCER recommendations, weighted according to the gravity of problems found (EMBRAPA)

Possible Information Sources:

Existing:	New:
• External Reviews • Center annual reports • Board of Trustee evaluations	• Employee Surveys • Self-reporting by centers

ELEMENT 7: Financial Health	
Description: The center's effectiveness in acquiring and managing financial resources.	
Potential Indicators: • Overall financial outcome for the year[9] • Liquidity (working capital)[9] • Fixed cost ratio[9] • Personnel expenditures as a % of total expenditures[9] • Ratio of projected to actual expenditures • % of funding that is unrestricted	
Possible Information Sources:	
Existing: • Center financial statements	New: • Self-reporting by centers

[9] Taken from indicators currently used by CGIAR

ELEMENT 8: Culture of Learning and Change

Description: The center's capacity to learn from its performance and environment and make appropriate changes to enhance performance

Potential Indicators:

- Mechanisms are in place to update strategies regularly.
- Mechanisms are in place to support and manage change.
- The center demonstrates a willingness to change.
- A system is in place to reward innovation.
- Internal evaluations are conducted and used to improve performance.
- % of budget spent on management and technical training.
- The center's staffing plans and profiles (e.g. healthy turnover rates, age profile, influx of "new blood").
- Mechanisms are in place to develop and deliver outputs by involving stakeholders.
- The extent to which the center gathers information relating to the present and future needs of stakeholders.

Possible Information Sources:

Existing:	New:
• External reviews • Center self-reports • Center personnel records • Reports on internal program reviews • Center mid-term plan	• Stakeholder surveys • Staff surveys