ASTI
Agricultural Science & Technology Indicators
FACILITATED BY IFPRI

FARA

THE ROLE OF EVALUATION IN STRENGTHENING AGRICULTURAL R&D IN SUB-SAHARAN AFRICA

Information, Instruments, and Actors

Leonard Oruko and Howard Elliott

Conference Working Paper 10

Prepared for the ASTI/IFPRI–FARA Conference | Accra, Ghana | December 5–7, 2011

AGRICULTURAL R&D: INVESTING IN AFRICA'S FUTURE
Analyzing Trends, Challenges, and Opportunities

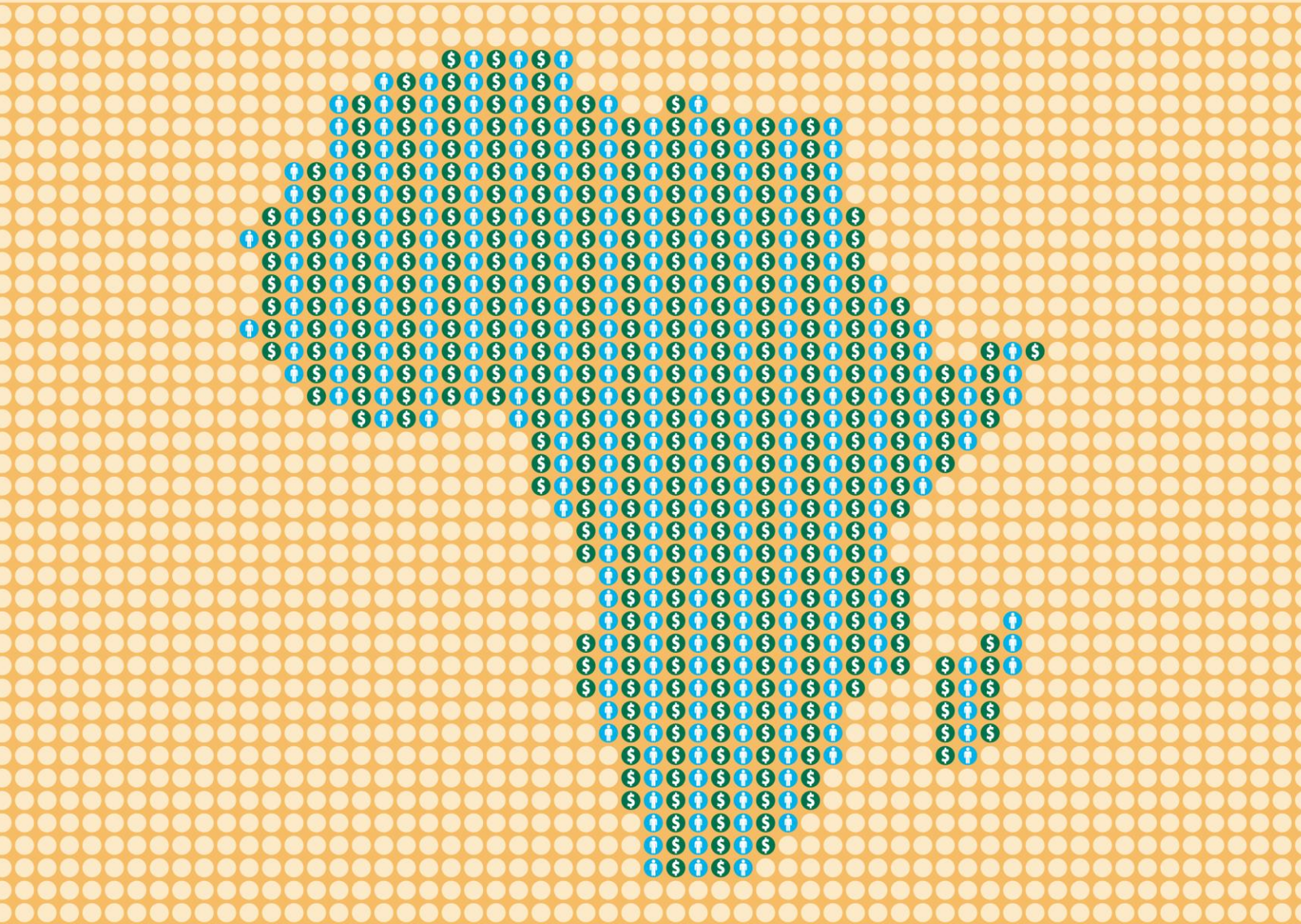

Table of Contents

List of Tables

List of Boxes

About the Authors

Leonard Oruko (loruko@yahoo.com) is an agricultural economist; **Howard Elliott** (howardelliott@shaw.ca) is an independent consultant based in Winnipeg, Canada, and the former deputy director general of the International Service for National Agricultural Research (ISNAR).

Acknowledgments

The authors gratefully acknowledge comments from John Lynam, the ASTI team, and an anonymous reviewer, as well as the excellent editing by Mary Jane Banks.

Acronyms and Abbreviations

ADII	Agricultural Development and Innovation Index
AIS	agricultural innovation system
AKIS	agricultural knowledge and information systems
ASARECA	Association for Strengthening Agricultural Research in Eastern and Central Africa
ASTI	Agricultural Science and Technology Indicators (initiative)
CAADP	Comprehensive Africa Agriculture Development Program
CCT(s)	conditional cash transfer(s)
CGIAR	Consultative Group on International Agricultural Research
CIMMYT	International Center for the Improvement of Wheat and Maize
CIP	International Potato Center
FTE(s)	full-time equivalent(s)
GFAR	Global Forum for Agricultural Research
IAR4D	integrated agricultural research for development
IARC(s)	international agricultural research center(s)
ICARDA	International Center for Agricultural Research in Dry Areas
ICRISAT	International Crops Research Institute for the Semi-Arid Tropics
IFARD	International Federation of Agricultural Research Systems for Development
ILAC	Institutional Learning and Change (CGIAR program)
INFORM	Information for Research Management (data capture tool)
ISNAR	International Service for National Agricultural Research
IITA	International Institute of Tropical Agriculture
INRM	integrated natural resource management
IRRI	International Rice Research Institute
KARI	Kenya Agricultural Research Institute
M&E	monitoring and evaluation
MDGs	Millennium Development Goals
NARI(s)	national agricultural research institute(s)
NARS(s)	national agricultural research system(s)
NARES(s)	national agricultural research and extension system(s)
NEPAD	New Partnership for Africa's Development
NRM	natural resource management
R&D	research and development
ReSAKSS	Regional Strategic Analysis and Knowledge Support System
SPIA	Standing Panel on Impact Assessment (of the CGIAR)
SROs	subregional organizations
SSA	Sub-Saharan Africa
UNEP	United Nations Development Programme
USAID	United States Agency for International Development

Abstract

This paper links the evolution of research monitoring and evaluation (M&E) tools to major questions being asked and to demands for research reform at the global, regional, subregional, and national levels. First, in order to inform decisionmaking, research must continuously push the frontiers, conceptually and computationally, while providing practical considerations to policymakers and their advisers. Second, the demand for new concepts and tools is continuously evolving as problems are redefined: each successful study enhances the demand for better data and better tools. Finally, monitoring paces the reform process in three critical ways: (1) it corrects pathways to new objectives as needed; (2) it measures the efficiency and effectiveness of implementation; and (3) it identifies organizational, institutional, and systemic constraints to achievement of objectives. This paper interprets several key reforms at national and international levels; relates them to the M&E tools used; and draws some conclusions for the future development of data, information, and analytical approaches in Sub-Saharan Africa.

1. INTRODUCTION

This paper purposely covers a wide canvas. It situates research evaluation in an evolutionary context that begins with an analysis of rates of return to investment in agricultural R&D and finishes with its contributions to addressing multiple objectives and the improvement of agricultural innovation systems in Sub-Saharan Africa (SSA). It has been necessary to develop both the tools and data to provide evidence for decisionmakers to put their support behind research. In the 1970s and 1980s, evidence on the high historical rates of return to commodity research underpinned the expansion of the Consultative Group on International Agricultural Research (CGIAR) and the consolidation of national agricultural research systems (NARSs). In spite of strong evidence, the case for increased investment was constrained by the context of structural adjustment and budget reform. As liberalization took place, donor concerns shifted to poverty and the environment, and new efforts were made to provide an economic analysis of the returns to natural resources management. With globalization, the potential for innovation from wider partnerships, cross-sectoral collaboration, and demand-driven research called for new insights into how agricultural research investment could be made more effective through a shift in focus from R&D to knowledge systems and innovation systems. This has changed many relationships and the focus of accountability.

Dealing with the "evaluation challenge," requires a consideration of many different aspects: data availability and maintenance, the client and the purpose of the evaluation, temporal and spatial attribution problems in measuring the impact of research through econometric approaches, impact pathways, and the design of experimental and quasi-experimental approaches to identify "what works" within controlled domains.

This paper introduce three projects that have created benchmark data to underpin policy analysis and, in some cases, decisionmaking: (1) The Agricultural Science and Technology Indicators (ASTI) initiative, which is facilitated by the International Food Policy Research Institute (IFPRI), (2) the Agricultural Development and Innovation Index (ADII), and (3) the Comprehensive Africa Agriculture Development Program (CAADP) monitoring and evaluation (M&E) system located at the Forum for Agricultural Research in Africa (FARA). The challenge is to go beyond measuring capacity to identifying ways of facilitating emergent systems, which by definition are self-organizing.

Considered are evaluation approaches that focus on learning and change in complex adaptive systems where impact pathways are imprecise and outcome mapping may be a more useful approach to evaluate success or failure. Examples are drawn from the evaluation history of eastern Africa at national and subregional levels, and that experience is related to FARA's role in monitoring and evaluating CAADP. Finally, the evaluation procedures developed by the CGIAR, which interface with systems at both national and regional levels, are examined.

2. THE EVOLUTION OF EVALUATION TOOLS: AN INTERPRETATION

Over the four decades from the 1970s onward, the global context has conditioned the demand for research. Concepts, tools, and information have evolved independently, helping research leaders to respond to changing political objectives and scientific challenges. This evolution is summarized in a highly schematic form, by decade, accompanied by a short narrative (Table 1). Some characterizations may overlap decades, but these do not distort the overall picture of interlinked policy, technology, and institutions.

Table 1. Evolution of context, research systems, and supporting evaluation tools

	1970s	1980s	1990s	2000s and beyond
Characterization of decade				
Economic context	• Instability	• Adjustment	• Liberalization	• Globalization
Policy goals	• Stabilize	• Get budget right	• Get prices right	• Get system right
Research paradigm	• Green revolution	• Farming systems research leading to natural resource management	• Poverty and environment	• Growth and Millennium Development Goals (MDGs)
Driver	• Science	• Policies	• Institutions	• Systems
Agenda	• Productivity enhancement	• Consolidation and capacity building of national agricultural research systems	• Agricultural knowledge and information systems, partnerships	• Agricultural innovation
Evolution of the Consultative Group on International Agricultural Research (CGIAR)				
Mission	• Availability of affordable food in tropical countries	• Sustainable food production	• International research and partnership with national agricultural research systems	• High-level science, partnerships, links to development in the South
Focus	• Productivity key cereals	• Resource conservation, productivity of production systems, policy environments, national research capacity	• Sustainable productivity increase, Nutrition and well-being , low-income people	• Program-based funding through consortium research projects; multicenter and multipartner research for development
Evolution	• New centers: livestock, ecoregional systems	• Addition resource-based centers	• Renewal and Lucerne Declaration: food security	• Performance contracts: centers–consortium; consortium-donor fund; fund council-donors
Evolution of national systems				
Structure and funding	• Departmental institutes, legacy infrastructure; government and donors	• National agricultural research institute consolidation and plans; International Development Association soft loans from the World Bank Group	• National agricultural research system master plans; recognition of agricultural knowledge and innovation systems; subregional organizations	• Innovation systems; regional and continental programs; funding through programs
Evolution of planning and evaluation requirements				
R&D Planning	• Departmental plans	• National agricultural research institute consolidated plan	• National agricultural research projects (multi-donor)	• National Agricultural Productivity Programs
Research evaluation	• Rate of return to commodity research	• Ex ante impact analysis (EAIA); priority setting	• EAIA with spillovers and distributional effects	• New tools to measure policy and natural resource management
Monitoring	• Logframe analysis	• Implementation against logframe	• Learning and change; participatory methods	• Performance measurement, results frameworks, identification of pathways, outcome mapping

Source: Compiled by authors.

The CGIAR was created in the 1960s and 1970s, and its expansion from the original four international centers (two commodity centers, the International Rice Research Institute [IRRI] and the International Center for the Improvement of Maize and Wheat [CIMMYT], and two ecoregional centers, the International Center for Tropical Research [CIAT] and the International Institute of Tropical Agriculture [IITA]) to include livestock and further eco-regional Centers (for example, the International Crops Research Institute for the Semi-Arid Tropics [ICRISAT] and International Center for Agricultural Research in Dry Areas[ICARDA]). The research paradigm was the Green Revolution, which was production-led with the later introduction of farming systems research and social sciences to NARSs through both commodity and eco-regional centers. Particularly in Africa, research in NARSs was still largely found in Ministerial departments and legacy institutes from the colonial period. Planning for research was within the relevant department using recently introduced tools, such as logframe analysis. Research evaluation was largely ex post rate of return analysis for commodity programs. The relative ease of dealing with commodities has carried on.

The 1980s was the decade of the NARSs, which saw the widespread consolidation of departmental and legacy institutes into national agricultural research institutes in Africa, and the move from national institutes in Latin America to "foundation models." The focus was on developing human and institutional capacity and consolidating research around national priorities. The CGIAR admitted a number of "resource-based" centers. As a period of structural adjustment ("get the budget right") there was restricted funding and concerns with internal sources of growth and stability. Tools for planning and budgeting for consolidated national programs were complemented by ex ante impact analysis and priority setting. A seminal publication by Alston, Norton, and Pardey (1995) was one output of work underway in the late-1980s. National agricultural research institutes (NARIs) were embracing "shortcut scoring methods" for priority setting (as described in the KARI case below).

The 1990s was a period of liberalization ("get the prices right"), but it also saw a moving away from the Washington Consensus to recognize that "institutions matter," and that they matter most for poor people. Concern institutionally was with the development of the "whole NARS" or "agricultural knowledge and information system" (AKIS). Scientists who worked originally on farming systems and natural resource management (NRM) broadened their perspective to integrated natural resource management (INRM) systems by incorporating policy and institutions. Subsequently, they turned their tools and attention from location-specific NRM systems to larger scales, embracing landscape planning and climate change. The CGIAR, or CG, system turned from the productivity of commodities and commodity-based production systems to addressing the reduction of poverty and protection of the environment. This was largely driven by the concerns of funding agencies, which looked to the CG to solve problems for which agricultural research was a blunt instrument. Subregional organizations (SROs), prominent in colonial days, were rediscovered as a means of collective action and achieving economies of scale among NARSs and their partners. Organizing collective action among independent countries has been more difficult than organizing regional centers under a single colonial control.

Since 2000, the era of globalization has waxed and waned. Ambitious growth rates were planned, and Millennium Development Goals (MDGs) were adopted. Getting national systems and multiple partners to work together, and facilitating the emergence of national agricultural innovation systems has changed the way research is planned, organized, and financed. The reform of the CGIAR is about the achievement of a set of system-level outcomes through center- and consortium-based research programs that depend on entirely new forms of partnership among centers, with a diverse set of actors.

3. EVALUATION CHALLENGES

It is useful, upfront, to introduce several challenges for evaluation. A first and basic challenge is the quality and availability of data. A second is clarity about the purpose of the evaluation and the

relationship between the evaluator and the entity being evaluated. Is the evaluation for external accountability, or is it utilization-focused to promote learning and change? A third challenge, deals with attribution: how the outcome or impact can be attributed in a causal way to an activity, factor, or event, and whether this relationship can be generalized. The fourth challenge is cost-effectiveness and the ability of the approach to answer the impact or attribution questions asked.

A basic challenge for evaluation is the quality of data available from public sources in meaningful time series that can be used to evaluate the return to investment. This includes output measures (production and productivity) and input measures, such as financial expenditure and human resources. Data quality is evaluated by its reliability, validity, and timeliness (Kusek and Rist 2004). The ASTI initiative created quality benchmark data and, to its credit, managed to maintain and improve the data over time. Maintenance of databases as international public goods has been a significant challenge for CG centers, and it is reassuring that the CG now recognizes the maintenance of data as one of its core responsibilities. As more of Africa is mapped for geo-referencing data by international programs such as the United Nations Environment Programme (UNEP), researchers are becoming adept at splicing data from diverse sources to create new databases. The definition of what constitute "data" may include the careful compilation of "cases" that provide context for identifying causes of success and failure that can be widely generalized. One clear suggestion is that failures need to be documented as diligently as successes (Jones 2009). While the statement "The plural of anecdote is data" is cited and debated in both a positive and a negative way, it reflects the fact that scientific investigation is often motivated by the desire to prove or disprove widely held beliefs that are based on some detectable distribution of personal experiences.

It is useful to briefly introduce three key evaluation issues that will be illustrated later by case studies and references to the literature.

Purpose. Is the evaluation primarily for accountability upward to funding authorities, or is it designed for learning and change by the program or institution being evaluated? Are both possible, and what independence is sacrificed? Is the funding of the evaluation shared appropriately by those who seek accountability, those who learn, and those who are objects of the evaluation?

Attribution. Does the approach allow the attribution of impact to factors, activities, or institutional variables? Does one work back from impact and seek to attribute outcomes causally or does one map outcomes as they occur and explain them by critical factors?

Choice of Approach. Does the chosen method answer the evaluation question and fit the context, and is it feasible with the resources available? The choice of approach is not only made on the basis of "rigor" of methodology, but also on the basis of cost, flexibility, and whether it defines a valid counterfactual: what would have been the situation in the absence of the project or intervention? A valid counterfactual can be generated through econometric techniques or through experimental and quasi-experimental design.

3. BENCHMARKING NATIONAL AGRICULTURAL RESEARCH SYSTEMS AND AGRICULTURAL INNOVATION SYSTEMS

It is useful at this point to follow up the discussion of the evolution of R&D systems and their evaluation challenges with a discussion of programs to benchmark that change: the aforementioned ASTI, ADII, and CAADP's M&E system. ASTI was initiated in the decade of NARSs. It focused specifically on establishing a database for tracking human and financial investments in agricultural research by research units aggregated to national institutes. By capturing disaggregated qualifications of human resources and full-time equivalent (FTE) researchers in universities, along with budgets it developed a picture of the "whole NARSs." ADII attempts to supplement published information about key actors and

processes in the broader national agricultural innovation system. CAADP's M&E system (ReSAKSS 2010) proposes to collect a core set of indicators for agricultural development at a much higher level of aggregation to monitor the continent's progress toward the achievement of CAADP targets. The database could also play an important role in monitoring regional integration.

The Agricultural Science and Technology Indicators (ASTI) Initiative

Research that creates data on research institutions, investments, and impacts has been an underappreciated part of the discipline. As Alston et al. (2009, 547 and 549) state:

> A significant part of the economic literature includes studies that describe, document, and quantify the institutions that fund, regulate, and conduct agricultural research as well as the investments they make. These "descriptive" studies are of value in their own right, but they also provide an institutional frame of reference and data for econometric and other modeling studies. . . . Compared with measures of productivity and its elements, measures of investment in research (and counterpart measures of stocks of scientific knowledge) have attracted much less effort and attention in the literature. This relative neglect could be comparatively pernicious. It takes a lot of work to develop measures of agricultural research investments. Appropriate measures of public agricultural research investments are not published in suitably long time series, in the relevant form, by any government agency.

For international data on agricultural research, ASTI is now approaching the 25th anniversary since its conception as the International Service for National Agricultural Research (ISNAR)–International Federation of Agricultural Research Systems for Development (IFARD)[1] Survey of National Agricultural Research Systems. The survival, and indeed the success, of ASTI is due to three factors: (1) the professionalism of its design, (2) the rigor with which data were collected and processed, and (3) the resistance of its leadership to pressure to collect data on every new fad that came along. ASTI was started with a strong theoretical grounding and clear hypotheses about research investments that it could be used to test. As a result, it became the international standard for data on human and financial investments in agricultural research. Its data are reliable, comparable over time and across countries, and regularly updated. Even if the frequency of updating has depended on the entrepreneurship of its leaders more than on the collective contribution of its users, it has been updated periodically. The data continues to be the basis for most analyses of the level and growth of funding, human resources, and gender capacity. The degree to which ASTI has succeeded in its niche is easily demonstrated by the use (and misuse) of two of its early findings that have been elevated to "memes": (1) the agricultural research intensity ratio (ARI) equal to or greater than one, and (2) the description of non–CGIAR investments as "the other 96 percent" of global investment are widely quoted if often incorrectly used. The first is often mis-specified as an investment target, and it has been decades since the CGIAR was 4 percent of worldwide investment in agricultural research. The lessons of ASTI are clear for investors and users: they must ensure adequate investment in this public good, and reinforce focus and professionalism on the part of its guardians. A long-term vision is essential.

Benchmarking of Agricultural Innovation Systems: Agriculture, Development, and Innovation Index

Spielman and Birner (2008) note that the introduction of an AKIS perspective in research planning fell short of what was needed: it treated agricultural research as one point in a "knowledge triangle" and not as part of a wider agricultural innovation system (AIS). Empirical studies to assess the entire AIS were scarce. They argued that the collection of data on innovation system inputs, processes, and

[1] Both ISNAR and IFARD are now defunct, but IFARD was a precursor of the current SROs.

outcomes is a necessary precondition for cross-country analyses that examine how different components of an innovation system and their relationships affect innovation performance in the agricultural sector. However, several issues arise in the construction of innovation indicators. First is the idea that innovativeness can be reduced to a single index value (much like GDP) for comparison across countries and over time. Second is the hypothesis that the relationship between innovation and the various inputs and processes identified as key determinants is not endogenous. Third is the fundamental issue of availability, both of the data in question and the resources to obtain data.

As with ASTI, the authors set out to select indicators with a strong theoretical grounding that explained their relationship to innovative performance (with respect to productivity, poverty reduction, and environmental sustainability). Spielman and Kelemework (2009b) attempt to provide a "proof of concept" that innovativeness in developing-country agriculture can be measured. The paper first identifies a set of indicators from secondary data sources that measure the key elements of an agricultural innovation system. Several hundred indicators are reviewed, validated, and aggregated into ADII. The paper then provides a toolkit for collecting and analyzing "systems-oriented" indicators that add more process-related nuances to ADII with both attributional and relational data. This is illustrated with data collected in Ethiopia and Vietnam in 2007–08.

The first step was to define the unit of analysis: three main subsectors of the agricultural economy (for example, for Ethiopia a main staple, maize; a high-value traditional export, coffee; and livestock, poultry). The second step was to diagnose the innovation subsystem. Using a network mapping tool (NetMap), they captured the formal and informal networks, relationships, and influences existing in each subsector. The third step was to collect a range of soft data through expert opinion surveys on organizational effectiveness, responsiveness to opportunities, and accountability to different stakeholders. The sample does not have to be large to get a good representation of the innovativeness of organizations in terms of efficiency-improving processes or improved products. The authors conclude that an expert survey approach can provide indicators of accountability, responsiveness, accessibility, effectiveness, and innovativeness that can significantly improve on secondary data used in the ADII. Ultimately, (their) paper "provides a combined qualitative/quantitative toolkit for measuring innovation systems properties and performance, and an analysis that emphasizes not only inputs and outputs, but also more difficult-to-measure systems properties" (Spielman and Kelemwork 2009b, 15).

The authors have made a good contribution to the literature in the IFPRI tradition of creating new value by merging data collected purposively by IFPRI with standard published databases. In this case they have merged quantitative and qualitative data to create a multi-criteria index of innovation system capacity that benchmarks systems for comparative purposes.

Ragasa (2010) demonstrates how one can work within an AIS framework to identify key factors explaining the productivity of individual researchers. Using a multi-level study of 300 randomly selected researchers, representatives of 47 public research institutions and universities in Nigeria, and 137 researchers of 16 agricultural research institutions and universities in Ghana, Ragasa shows that organizational climate (facilities, Internet access, leadership, work environment) play a much more important role relative to salaries and wages than economists might think (for more information, see Ragasa 2011). She also raises important questions about the role of gender in leadership.

The Comprehensive Africa Agriculture Development Programme

CAADP is an initiative of African Union's New Partnership for Africa's Development (NEPAD). As a framework for guiding coordinated and harmonized investment in agriculture, CAADP aims to help African countries achieve agriculture-led development that reduces mass poverty, food insecurity, and hunger, in line with MDG 1, halving poverty and hunger by 2015. Recognizing the need for an objective approach to measuring progress toward this goal and the achievement of results, CAADP's M&E framework identifies specific indicators on processes put in place; commitments and investments made; the agricultural sector; and changes in poverty, hunger, and food and nutrition security. More

importantly, a clear analytical agenda informed by ex ante impact evaluation, monitoring and operational evaluation and an ex-post impact evaluation generates an annual trends and outlook report on key performance indicators. Supported by IFPRI, the Regional Strategic Analysis and Knowledge Support System (ReSAKSS) initiative is mandated to implement CAADP's M&E framework.

A number of lessons can be drawn from the ReSAKSS experience over the past three years. Drawing on IFPRI's tradition of rigorous analysis, the ex ante evaluation of investment options employs a variety of simulation models depending on availability of data and information to predict growth and poverty reduction outcomes (Benin et al. 2008a, 2008b, and 2008c; Breisinger et al. 2008; Diao et al. 2008; Johnson et al. 2008; Thurlow 2008; Fan et al. 2009). By the same token, analysis of past investments in agriculture across and within countries reveals clear trends and outcomes of policy choices and investment options (Benin et al. 2010; Chilonda, Machethe, and Minde 2007; Chirwa et al. 2008; Govereh et al. 2009; Njiwa et al. 2008). From a typical M&E standpoint, these analyses not only provide more in-depth baseline data and information, but also a clear rational for setting targets based on projections from simulations and evidence of past performance.

The strategy of infusing analytical rigor in tracking the output, outcome, and impact indicators has, however, come at a price. As indicated earlier, the outputs from these analytical pieces are only as good as the quality of data. Accordingly, the choice of the standard indicators that are measured and reported on a regular basis was informed by the availability of data and information at the country level. This state of affairs has somewhat restricted the scope and depth of analysis to areas where data is readily available. To date, therefore, while the regular trends and outlook reports provide good information for advocacy at the strategic level, they do not inform adaptive learning for performance improvement in the short run at the operational level. Needless to say, the expectation that the annual trends and outlook report from ReSAKSS is the only source of all performance information to guide lessons learned at the program implementation level is unrealistic. The CAADP implementation strategy is organized around four pillars:

1. extending the area under sustainable land management and reliable water control systems,
2. improving rural infrastructure and trade-related capacity for market access,
3. increasing food supplies and reducing hunger, and
4. promoting agricultural research and technology dissemination and adoption.

Each of these pillars has a framework document with clear deliverables. It is therefore more realistic for each pillar to develop an M&E plan that provides more detailed evidence of performance—for example, how targeted investments in rural infrastructure lead to improved market access and increased incomes for smallholder farmers. Of more relevance to the subject under review in this paper is how strategies and programs for coordinating investment in agricultural research and development (R&D) would lead to improved productivity, income growth, and poverty reduction outcomes. The following section reviews approaches to evaluating research and development initiatives.

4. THE ROLE OF EVALUATION IN IMPROVING R&D: SOME FUNCTIONS AND TOOLS

To recap, so far in this paper the question of the focus of evaluation over time has been linked to both the economic context and the prevailing R&D paradigm. At the start of the Green Revolution, IRRI's response to famine in Asia was to "increase the pile of rice" (a powerful metric). Later rate of return analyses showed that the investment was repaid many times. Economic surplus analysis showed that increased productivity had led to reduced prices for consumers, while increasing net returns to farmers. The success of R&D gave following generations the chance to explore whether the technology favored large versus small farmers, men versus women (equity), and present versus future generations (sustainability).

In recent years, researchers have developed many ways to measure poverty (and its reduction) and relate this to investment in R&D. The inclusion of food security and nutrition concerns to outcomes at the CGIAR system level will accelerate research to develop indicators as well. The term "evaluation" is used here in its broad sense to include a range of techniques from ex ante impact assessmentswhich supports priority setting; through monitoring and performance assessment, which tracks the implementation of projects and achievements; to ex post impact assessment which measures the attainment of outcomes.

An Assessment Continuum with a Menu of Approaches

Many types of evaluation are needed: they serve different purposes, address different challenges, and answer complementary questions. This section addresses the contribution each makes to the "reform of R&D." Ex ante impact assessment is designed to answer whether an investment is likely to have a positive return, and, if there is a shortage of investment funds, provide information for choosing among alternative investments. (If you don't know where you are going, getting there faster doesn't help.) Ex post impact assessments is designed to provide information on returns to investment. In order to instill some consistent terminology, definitions from the most recent U.S. Agency for International Development (USAID) evaluation policy (USAID 2011) are used (Box 1).

Box 1. Concepts and consistent terminology

To ensure consistency in the use of key concepts, the terms and classifications highlighted below are used by USAID staff and those engaged in USAID evaluations.

- Evaluation is the systematic collection and analysis of information about the characteristics and outcomes of programs and projects as a basis for judgments, to improve effectiveness, and/or inform decisions about current and future programming. Evaluation is distinct from assessment, which may be designed to examine country or sector context to inform project design, or an informal review of projects.
- Impact evaluations measure the change in a development outcome that is attributable to a defined intervention; impact evaluations are based on models of cause and effect, and require a credible and rigorously defined counterfactual to control for factors other than the intervention that might account for the observed change. Impact evaluations in which comparisons are made between beneficiaries that are randomly assigned to either a treatment or a control group provide the strongest evidence of a relationship between the intervention under study and the outcome measured.
- Performance evaluations focus on descriptive and normative questions: what a particular project or program has achieved (either at an intermediate point in execution or at the conclusion of an implementation period); how it is being implemented; how it is perceived and valued; whether expected results are occurring; and other questions that are pertinent to program design, management, and operational decisionmaking. Performance evaluations often incorporate before–after comparisons, but generally lack a rigorously defined counterfactual.
- Performance monitoring of changes in performance indicators reveals whether desired results are occurring and whether implementation is on track. In general, the results measured are the direct and near-term consequences of project activities.
- Performance indicators measure a particular characteristic or dimension of project results (outputs or outcomes) based on a project's results framework and underlying theory of change. In general, outputs are directly attributable to the program activities, while project outcomes represent results to which a given program contributes, but for which it is not solely responsible.
- Performance management (managing-for-results) is the systematic process of monitoring the achievements of program activities; collecting and analyzing performance information to track progress toward planned results; using performance information and evaluations to influence decisionmaking and resource allocation; and communicating results to advance organizational learning and communicate results to stakeholders.

Source: USAID 2011, 1.

As the USAID definitions show, between ex ante appraisal and ex post impact analysis, a number of performance measurement functions exist that are normally assigned to operational managers. They are designed to provide accountability for the use of inputs and implementation of activities believed to lead to outputs and outcomes along an impact pathway. In the business literature, an approach like the "balanced scorecard" connects every unit of an organization to the organization's mission through key performance indicators in four dimensions: (1) customers, (2) finances, (3) internal business processes, and 4) learning and growth (Kaplan and Norton 1996). The approach has been adapted for use by governments and nonprofits by Niven (2003). It is an approach that could be useful to the CG as it tries to bring center and consortium research programs together around system-level outcomes.

The Evaluation Gap Working Group (CGD 2006, 2) argued that "rigorous" impact assessment is underdeveloped relative to the other types of evaluation.

> . . . Governments spend substantial sums on evaluations that are useful for monitoring and operational assessments, but do not put sufficient resources into the kinds of studies needed to judge which interventions work under given conditions, what differences they make, and at what cost. Impact evaluations complement other studies. Critics sometimes claim that impact evaluations can only tell whether something has an impact, not why and how. But a good impact evaluation can provide evidence about the mechanism through which the outcome is achieved when it simultaneously collects information on processes and intermediate outcomes. Impact evaluations are not a replacement for sound theories and models, needs assessments, monitoring and operational evaluations.

Incentives to do rigorous evaluation may be weak for several reasons. First, operational managers are reluctant to use scarce budget and time to commission an evaluation for the purpose of building USAID's knowledge base about what works, what doesn't, and why (Blue, Clapp-Wincek, and Benner 2009). Second, the pressure is for researchers to deliver results too quickly and for policymakers "there is still a pronounced hunger for success stories but a tendency to choke on failure" (Ruth Levine quoted by Chapoy 2011, 1). There may be a conflict of interest between the entity that demands the evaluation and the group that pays when tests of increasingly expensive evaluation methods are built into development programs. It is difficult to determine an excessive evaluation cost. How much should an agency spend on monitoring and evaluation? (The estimate of 3–5 percent by Blue, Clapp-Wincek, and Benner 2009 was anecdotal and taken from an unrelated sector.)

5. IDENTIFYING CAUSALITY

Two impact evaluation approaches seem to be competing: ex post impact analysis rooted in impact pathways, and experimental or quasi-experimental analysis rooted in randomized controlled trials. Both methods accept the principal that evaluation requires a comparison between the situation *occurring with the project* and a valid *counterfactual* describing the situation that would have occurred in the absence of the project. Both methods recognize the need to engage policymakers at the beginning of the process and both admit that documenting failures as well as success is important. Finding out what does not work is also important. As documented by Rasheed, Hall, and Vmsidar Reddy (2011), many lessons from the Research into Use projects of the Overseas Development Institute have not been applied. Finally, both approaches claim to accommodate—even gain validity—by more "ethnographic" analysis of what works and why it works. Table 2 attempts to lay out the advantages of each approach, suggesting where one might be chosen over the other.

The CGIAR guidelines for ex post impact assessment (CGIAR Science Council 2008a) express four unifying themes of good evaluation: (1) the need for a sequential approach,[2] (2) the need to go further along the input–output–outcome–impact pathway, (3) the need for better preparation of the ex post impact assessment based on elaboration of impact pathways, and (4) the need for transparency and analytical rigor (including an explicit counterfactual). The Science Council underlined that ex post impact evaluation, defined in this way, serves both accountability and learning purposes.

Table 2. A comparison of ex post impact assessment and quasi experimental studies

Comparison	Ex post impact analysis (CGIAR)	Quasi-experimental study
Purpose	Upward accountability; validation of impact pathways	Upward accountability and legitimation; identification of causal factors in outcome
Approach	Two-stage (sequential approach): Stage 1: focus on outcomes and impacts linked to adoption Stage 2: focus on impacts further along the impact pathway (for example, poverty reduction, food security)	Randomized controlled trials Large sampling frame
Flexibility in attribution of impact; adjustment to emergent situations	Can work backwards on impact pathway to attribute impact to responsible factors, or forward from activities through outcome mapping	Once the experimental design is in place it is difficult to modify if the initial conditions change or nonlinear responses occur
Scalability	Understanding of causal factors widely applicable	Limited if population is heterogeneous and opportunity for replication is small
Recommended use and cost	Generally applicable. Cost budgeted within project budgets (approximately 1–2 percent of total cost	Expense is justified where program is innovative, untested, and strategically relevant or major investments depend on rigorous evaluation

Source: Compiled by authors.

The call for rigor does not automatically prescribe quantification over qualitative methods although the Standing Panel on Impact Assessment expressed its preference as follows in its forward to an impact assessment of agricultural research in South Asia (Hazell 2008, ix):

> While SPIA fully concurs with the need to develop robust and relevant social and environmental indicators of the impact of agricultural research, it considers measurable indicators preferable to qualitative ones, as they lend themselves to wider application and aggregation of effects when scaling up.

The cost of evaluation for the new CGIAR, based on comparable organizations, is estimated at roughly 2 percent in the draft evaluation policy.[3] Of this, 1 percent would go to a central evaluation function, and 1 percent would be budgeted within the consortium research projects (CGIAR 2011a). Turning to what its proponents call "high-quality evaluation," Gertler et al. (2011) provide a useful rationale and checklist of criteria for when it makes sense. To justify mobilizing the technical and financial resources needed to carry out a high-quality impact evaluation, the program to be evaluated should be:

1. *Innovative*. It is testing a new, promising approach

[2] The guidelines describe two main types of ex post impact assessment: *economic rate-or-return assessments* and *multidimensional impact assessments.* The former focus on what are called Stage 1 outcomes and impacts (those that are closely linked to adoption, such as farmers' incomes), while the latter relate to Stage II impacts (those that are evident further along the impact pathway, such as poverty reduction and food security). The guidelines recommend that Stage 1 ex post impact evaluations should be rigorously carried out before other consequences along the impact pathway are evaluated.

[3] IDRC estimates central evaluation expenditure at 1.5–2.0 percent of total expenditure. FAO has established in its basic texts that 0.8 percent of total regular budget (including the administrative budget) is to be devoted to evaluation and 1.0 percent of the budget from nonregular budget resources which are principally for various forms of technical cooperation.

2. *Replicable*. The program can be scaled up or can be applied in a different setting
3. *Strategically relevant*. The program is a flagship initiative that requires substantial resources, covers (or could be expanded to cover) a large number of people, or could generate substantial savings
4. *Untested*. Little is known about the effectiveness of the program, globally or in a particular context
5. *Influential.* The results will be used to inform key policy decisions

In exaggerated form, a $1 million evaluation of a pilot project costing only $500,000 could be justified if it provided rigorous justification for scaling up a project costing several billions of dollars. Table 3 summarizes three illustrative cases where the rigorous design of the original experiment and the rigor of the evaluation were determinant factors in the degree of acceptance of the results and uptake by policymakers and end users.

Table 3. Three cases: Rigor of project design, evaluation of outcome, and project follow-on

Case	Purpose	Evaluation	Outcome
Millennium Villages (Nigeria, Ghana, Malawi)	Experimental intensive package intervention to spark sustained local development in Africa	Rigorous impact evaluation not possible due to weakness in project design: choice of sites, small number of comparators, choice of primary indicator; benefits based on project's "before and after" comparison were reduced when compared with trends in surrounding areas.	Suggestions for design of future evaluation made with small increase in marginal cost; Major investments require rigorous proof of concept: net gain and sustainability
HarvestPlus/International Potato Center (CIP): Orange-Fleshed Sweetpotato (OFSP) project (Mozambique, Uganda)	Increase vitamin A intake and serum retinol concentration in small children	Successful proof of concept: (1) improve beta carotene in OFSP; (2) retain pro-vitamin A after processing, (3) increase nutrition knowledge and demand for OFSP, (4) some change in production and marketing	CIP obtained support to work on (1) advanced breeding in Africa and (2) impact pathways for nutrition. Second stage on OFSP production and marketing systems: (1) increase demand, (2) develop commercial vine production and (3) improve market access through value chain development; randomized controlled trials are potentially applicable to testing strategies for reaching end users
Progresa (Mexico)	Anti-poverty and human resource investment conditional cash transfers (CCTs)	Maintenance of CCTs due to persuasive and credible evaluation of well-designed experiment proving CCTs impacted health and school leaving positively	Program continued, expanded, and model spilled over to other countries; prestige of Mexican academics enhanced by success; IFPRI contribution to policy environment and dissemination

Sources: Data on Millennium Villages are from Clemens and Demombynes (2010); data on the Orange-Fleshed Sweetpotato project are from Low et al. (2009 and de Brauw (2011); and data on Progresa are from CGIAR Science Council (2008b).

In the case of the Millennium Villages, critics contested the estimate of gains from a "before and after" study with no valid counterfactual (Clemens and Demombynes 2010). The Center for Global Development review of the data from their mid-term evaluation lowers the magnitude of the estimated "before and after" benefits by making comparisons with trends in the broader community. While the

"broader community" is not a true "control" group it was more representative of a counterfactual than the "before and after" estimate in the target communities. Precisely because the Millennium Village model was highly publicized and called for expansion throughout Africa before any proof that it was capable of delivering reduced poverty and sustained economic development, it should be held to a high level of proof. Its principal indicator (mortality of children under five years) was not a proxy for all the benefits promised, and it was not possible to say which parts of the package were essential. Clemens and Demombynes (2010) demonstrate that a more rigorous design of the project could have been done at small marginal cost, and they show how data from the current project design could create a counterfactual for the next review. Applying Gerter's criteria for the choice of evaluation technique, the Millennium Village project requires high-quality review because it was strongly publicized, claimed replicability, and would have been strategically relevant to the achievement of MDGs if implemented at scale.

In the case of the HarvestPlus/International Potato Center (CIP) Orange-Fleshed Sweetpotato (OFSP) project, the rigor of the agricultural and nutritional trials established several necessary facts. First, the breeders established that there was enough heterosis in sweetpotato to permit breeding to raise beta carotene to levels that could address vitamin A deficiency. Second, nutritional tests showed the bioavailability of pro-vitamin A was retained after both traditional and modern forms of processing. Third, full clinical trials with blood testing showed that modest consumption of OFSP raised serum retinol levels in children sufficient to prevent Vitamin A blindness. Testing of knowledge and awareness showed that women in the project areas had better knowledge of the nutritional advantage of OFSP, but this seemed to have limited impact on nutritional habits. This finding led CIP to address the problem of reaching end users in a systematic way by increasing demand in the target population and developing institutional markets. It also addressed the need to maintain the quality of planting material through commercial vine production. The demand for commercial vine production could only come about through improved value chains. The rigor of the nutritional work plus economic analysis showed that a food-based approach to vitamin A deficiency could be competitive with supplementation in reaching the critical preschool population. The approach proved a concept, leading to programs to scale out the approach in three regions. The program is integrated from building advanced breeding capacity in each region through market access and demand creation (SASHA 2009).

In the case of Progresa, the superiority of the conditional cash transfer (CCT) program over other social programs *in improving health and retaining students in school* was demonstrated through a rigorous evaluation using randomized control trials. The indicators were unambiguous, and the experience across comparators was clear. The rigorous evaluation is credited with maintaining the program and providing proof of concept that led to its rapid spillover to other countries and regions. The case is widely cited because it meets all six of the World Bank criteria mentioned above for investment in "high-quality evaluation." The evaluation itself, when reviewed as an exercise in "policy research," highlighted a number of public-good benefits that transferred a low cost: (1) the proof of concept entered the policy discussion widely; (2) the Mexican academics involved in the evaluation gained prestige and access to policymakers on other issues; and (3) IFPRI's role was present, although indirect, through methodology in the public domain, widening of the policy debate, and dissemination of findings. The issue of "attribution of impact" is a difficult one, as recognized in this example, since IFPRI was neither directly involved in the evaluation, nor in advising the Mexican government at the time (CGIAR Science Council 2008b).

The take-home message from these studies is that *rigorous evaluation has longevity and long legs.* The studies are cited for many years after publication and the findings enter policy debates beyond local borders. Duflo (2011) argues that policymaking is a clear prerequisite for program design, but it is often subject to "lazy thinking" and the "three-I problem: ideology, ignorance, and inertia." Her research and that of others has demonstrated that an increasing number of policy and social issues lend themselves to randomized evaluations or controlled field experiments (Duflo 2011; List 2011; Gertler et

al. 2011). Policy is a particularly important area for study because good policy, often at decentralized levels, can eventually result in good politics even in countries with bad *institutions*—meaning national customs, laws, and norms of behavior (Duflo 2011).

6. EVALUATION FOR INSTITUTIONAL LEARNING AND CHANGE: DEALING WITH COMPLEXITY

Created in 2003, the Institutional Learning and Change (ILAC) program of the CGIAR generates new knowledge on innovation processes through applied research and evaluation and strengthens partnerships to foster pro-poor innovation. With creation of the Consortium, ILAC's stated objective is to "support the efforts of the CGIAR Consortium and Centers, the Fund Council, and other stakeholders to steer the change process." It proposes to do this through action-research, leadership building, and exchange of experiences (Ekboir and Sette 2011, 4).

ILAC argues that "rigorous determination of causality is not defined by a particular method, quantitative or qualitative, but by clearly stating the assumption and logic used" (Ekboir 2011, 28). Which approach is more appropriate? It depends on the goals and the problem under study:

- quantitative approaches are useful for the study of relatively stable, simple relationships that hold for large numbers of cases;
- qualitative approaches are appropriate for the analysis of complex relationships that change over time or space; and
- quantitative methods are less effective (that is, less rigorous) in complex processes because they limit exploration of possible explanations and bet that the posited explanation is the closest to the truth.

In short, Ekboir and Sette (2011, 29) conclude that "for accountability, simple quantitative methods are probably better than sophisticated models and qualitative analysis; for learning, qualitative methods supported by simple quantitative indicators are definitely better." ILAC has been historically prominent in promoting the use of outcome mapping (Smutylo 2005) and participatory impact pathway analysis (Douthwaite et al. 2008). It worked with FARA in 2005 on mainstreaming institutional learning and change in the Sub-Saharan Africa Challenge Program (Acosta, Jones, and von Kaufman 2005).

The ILAC program was an early proponent of "utilization-focused evaluation," which centered on the beneficiary being able to use the evaluation to plan and implement change (Patton 2008; Patton and Horton 2009). More recently, Patton (2010) has introduced the concept of "developmental evaluation" in which the evaluator is internal to the innovation process. In this approach, "pre-formative" evaluation plays a primary role and summative evaluation may not be possible because of the many changes that have taken place.[4]

> Developmental evaluation supports innovation *development* to guide adaptation to emergent and dynamic realities in complex environments. A complex system is characterized by a large number of interactions and interdependent elements in which there is no central control; self-organizing and emergent behaviors based on sophisticated information processing generate learning, evolution, and development. . . . Informed by systems thinking and sensitive to complex nonlinear dynamics, developmental evaluation supports social innovation and adaptive management. Evaluation processes include asking evaluative questions, applying evaluation logic, and gathering real-time data to inform ongoing decision making and adaptations. The evaluator is often part of a development team whose members collaborate to conceptualize, design, and test new approaches in a long-term ongoing

[4] Formative evaluation focuses on improving a model; summative evaluation leads to a conclusion about whether the program was effective; developmental evaluation has the purpose of helping develop an innovation, intervention, or program.

process of continuous development, adaptation, and experimentation, keenly sensitive to unintended results and side effects.

Developmental evaluation (Patton 2010, 1) may be describing the potential role of ILAC as a system facility in supporting the Consortium while the CGIAR system maintains an independent evaluation arrangement for external accountability.

Evaluation and Adaptive Management in the CGIAR

The key characteristics of the monitoring system in the new CGIAR were described before the strategic results framework had been finalized.[5] However, it is agreed that the Consortium and CGIAR Fund donors are "mutually accountable" for Consortium research project (CRP) outcomes. The monitoring system regarding research under the Strategic Results Framework is the overall responsibility of the Consortium, which is also responsible for commissioning evaluations of the CRPs. CGIAR accountability resides in four main areas: (1) strategic impact, (2) quality and relevance of program performance, (3) managerial and governance performance, and (4) financial performance and resource mobilization.

Developing an evaluation system in the "New CGIAR" still presents some key challenges. First, there is the operational urgency to develop an accurate and harmonized performance monitoring system for Centers and CRPs. Accountability for system-level outcomes is a complex web of mutual accountability relationships. Administratively, a system is yet to be developed. Performance contracts will be different with different partners inside and outside the system. A technical fix of a Web-based portal is not a silver bullet to achieve a degree of harmonization that has not been possible heretofore. At the impact end of the evaluation process, the attribution of impact for multiple outcomes dependent on joint contributions is not yet resolved.

Using adaptive management approaches, ILAC's activities in 2011 and 2012 will be organized along three lines of work:

1. helping to develop new types of partnerships, devise indicators for monitoring the evolution and management of these partnerships, and explore new incentives to increase the effectiveness of researcher participation in innovation networks;
2. creating a space for reflection where CGIAR stakeholders and experts can think collectively about the dynamics of poverty and agriculture the niches the CGIAR should occupy, the types of partnerships it should engage in, and the type of science it should conduct; and
3. contributing to the change process by (a) drawing lessons from experiences of organizational change, and (b) participating in change processes in the Consortium and individual Centers—for example, assessing methodologies to foster organizational change and facilitating access to specialized information on organizational change.

Since activities will be increasingly carried out by CRPs, in which strategic alliances, partnerships, and joint ventures will be the modality, the "attribution problem" will become more acute. Simple attribution of impact based on staff time, expenditure, or origin of parental lines will no longer be acceptable. ILAC introduced the notion of "contribution analysis" as a way of approaching this issue (Mayne 2008).

Lessons from the Overview: A Recap

The above overview of research evaluation has been selective but designed to point the reader to key issues and sources. For the authors, it has demonstrated that the discipline is evolving to meet demands for better prioritization, project design, accountability, and adaptive management.

[5] Based on October 26, 2009, draft "M&E Framework for the New CGIAR."

- Serious analysis of the policy environment and context are necessary to avoid the three "I" problems of "ideology, ignorance and inertia." The investment in such policy research could prevent much waste. Good policies could even lead to good politics (Duflo 2011).
- The marginal cost of designing a rigorous evaluation framework is small compared with the cost of doing it poorly and living with mistakes (Clemens and Demombynes 2010).
- High-quality evaluation is called for when a program is innovative, replicable, strategically relevant, untested, or influential (Gertler et al. 2011). Research evaluation is an investment, and allocation of costs needs to be subject to norms (Blue, Clapp-Wineck, and Benner 2009). The incentive to pay for evaluation may be related to the degree to which it is utilization-focused (Patton 2008).
- High-quality evaluation may be both quantitative and qualitative (Ekboir 2011). Quantitative methods are preferred by SPIA because they can be aggregated and compared more easily (Hazell 2008)
- Evaluation should be rigorously built into programs to establish a base for performance-based M&E.
- Descriptive studies and data collection are invaluable activities that are underfunded (Alston et al. 2009). It is necessary to mine evaluation studies for lessons; many are never put into practice (Rasheed, Hall, and Vmsidar Reddy 2011).
- Benchmarking databases need to protect their salience, credibility, and legitimacy. They need to be appropriate for the purpose and not lose focus (for example, ASTI and ADII), and they can evolve to meet new needs without losing focus. Every attempt to measure a phenomenon creates a demand for better information and leads to refinement of concepts.

Capacity building for evidence-based policy advice at the national level is necessary: ideas travel fast, data can feed debate, and analysis can legitimize choices. National experts do this best (for example, Progresa).

7. MONITORING AND EVALUATION EXPERIENCES IN AFRICA AT THE CONTINENTAL, SUBREGIONAL, AND NATIONAL LEVELS

Over the years, numerous attempts by African agricultural R&D organizations to establish functional M&E systems have generated mixed results. It is therefore difficult to find a fully functioning M&E system in a NARI at this time. Earlier in this paper M&E was framed according to three functional categories: ex ante impact evaluation, monitoring and operational assessments, and ex post impact evaluation.

The agenda to institutionalize operational M&E and impact assessment in African NARIs was largely spearheaded by the ISNAR in the late-1980s and early 1990s. With the support of ISNAR, the Kenya government conducted a root and branch review of its agricultural research system in 1986. An autonomous body, the Kenya Agricultural Research Institute (KARI), was created thereafter with an administrative headquarters and a network of national and regional research centers.

Monitoring, Evaluation, and Impact Assessment at the Kenya Agricultural Research Institute

In keeping with ISNAR's mandate to strengthen research management in NARSs, KARI—then considered as one of the strongest NARSs in Africa—received technical support to conduct priority setting for its programs. Employing the subjective scoring method, a "Blue Book" summarizing national and regional priorities (within Kenya) was produced. Significantly, the Blue Book was prepared by Kenyans and totally owned by KARI. In spite of some errors in the application of short-cut scoring methods, it was a major step forward and built both confidence and commitment for the next effort.

The Blue Book was followed by a methodological jump forward to a more objective economic surplus approach to ex ante impact evaluation in the 1990s (Mills 1998). The role played by the Director of KARI as a strong champion of this effort was essential. As a spinoff, the output of the exercise became one of the most cited ISNAR (with KARI) publications, and a reference for other countries.

Another strong influence from international centers was manifested in seminal studies to estimate the rates of return to research investment in the maize program (Karanja 1990. These developments suggest that KARI began the process of institutionalizing impact evaluation way back in the late-1980s. Moreover, it employed a combination of spatial and economic analyses to prioritize interventions in maize research and NRM. This identified the need for a new variety of maize to serve a key recommendation domain that ultimately became a Kenyan success (Hassan 1998). The objective approach to priority setting was almost institutionalized in the 1990s; however, retention of planning skills and perhaps the level of enthusiasm was not sustained.[6] At the present time, KARI employs more of the subjective scoring approach to priority setting.

Along with these approaches, efforts to strengthen the operational M&E function were also driven by ISNAR. Through a series of training exercises, ranging from logframe analysis to data capture using the Information for Research Management (INFORM) data capture tool, KARI scientists and managers were sensitized to the need to generate timely performance information. For several reasons, these efforts were not sustained, and KARI drifted back to the business-as-usual way of tracking and reporting research implementation through yearly technical reports. First, the INFORM program was suitable for collecting program planning information and linking it to human resources, but it lacked a link to the accounting system (this was before it was easy to have systems communicate with each other). As such it was an incomplete management information system that would not link to ex post evaluation (although this was not yet contemplated). Successive efforts to bring in expensive commercial systems ran into problems, and momentum was lost. (The commercial systems did not handle the program planning well and required expensive add-on modules). Bottomline, a management information system that can be modified by one's own experts is a functional a necessity.

Monitoring, Evaluation, and Impact Assessment at the Association for Strengthening Agricultural Research in Eastern and Central Africa

The Association for Strengthening Agricultural Research in Eastern and Central Africa (ASARECA) is a nonpolitical SRO serving the national agricultural research and extension systems (NARESs) of 10 countries: Burundi, the Democratic Republic of Congo, Eritrea, Ethiopia, Kenya, Madagascar, Rwanda, Sudan, Tanzania, and Uganda. ASARECA came into being primarily as a platform to redress market failure in the provision of agricultural R&D in Eastern and Central Africa. This was to be achieved by rationalizing research efforts to generate economies of scope and scale, and promoting spillovers. With support from USAID to strengthen its role as a coordinating agency, ASARECA institutionalized performance monitoring as a management function to track implementation of its programs and to see that the original planned results of its networks and programs were achieved.

In 2005, ASARECA contracted IFPRI to conduct a strategic analysis and identify priority areas for investment (Omamo et al. 2006). Employing an innovative approach combining spatial, economic, and institutional analysis, the ex ante evaluation identified investment options that would have the greatest impact on poverty by generating broad-based agricultural growth. To date, ASARECA's strategic planning at all levels is informed by this analytical piece. With the merging of 16 ASARECA networks, programs and projects (predominantly CGIAR–managed networks and ASARECA–managed projects and programs) into 7 ASARECA–based programs, operational plans for each program were to be developed with a

[6] Many of the authors have gone on to international careers where their experience has been used. Some of them transited SROs before moving to international organizations, donor agencies, and foundations.

logframe base. With these major efforts going into planning, ex post impact assessment at ASARECA has not yet been developed, and some programs have yet to complete their strategic plans.

Monitoring, Evaluation, and Impact Assessment at the Forum for Agricultural Research in Africa

FARA was created in 2001 as a facilitating and information exchange forum to support the SROs. In the subsequent years, FARA has evolved into a continentwide umbrella organization for agriculture, bringing together and forming coalitions of major stakeholders in agricultural research, extension, education, and training. FARA's strategic objective is to contribute to broad-based agricultural productivity growth in Africa by (1) supporting the development of effective R&D institutions; (2) enhancing access to agricultural technologies, knowledge, and information; (3) advocating for supportive policies and institutions; (4) building capacity for agricultural innovation; and (5) supporting agricultural innovation platforms.

In 2010, FARA developed an M&E strategy for measuring and providing evidence of performance. Based on a clear impact pathway, tracking and reporting achievement of results primarily serves to inform implementation and operational lesson learning. In addition, the strategy identifies ex post impact assessment as a priority activity in order to measure progress toward the achievement of growth and poverty reduction goals.

The importance of measuring and documenting the impact of investments in agricultural research and development was recognized long before the development of FARA's M&E strategy. The Sub-Saharan Africa Challenge Program, one of FARA's flagship programs, was approved by the CGIAR in 2004. It put forward the concept of a new paradigm: integrated agricultural research for development (IAR4D) and called for proposals to develop benchmark sites in three regions of Africa where a range of themes could be tested. Three sites were accepted through quite different processes in the three regions. The CGIAR Science Council questioned whether there was sufficient scientific content to IAR4D as a construct, but did not hold up the FARA–owned program. Given its slow start, there was inconclusive evidence of impact by the time of the first external review in 2006. Following the evaluation the Challenge Program was redesigned to include an ex post impact evaluation using randomized controlled trials to test the IAR4D concept. FARA was thus one of the first African organizations to employ quasi-experimental approaches to test whether IAR4D would produce regional and international public goods (a CG concern) produce better impact under African smallholder conditions than traditionally used research and extension approaches. It was an opportunity to contribute to the debate in the CGIAR on NRM research, where international public goods were defined in terms of methodologies, new knowledge, and decision support systems. Since NRM research was often very location specific, it was also an opportunity to test whether IAR4D could be scaled up.

The Second external evaluation (CGIAR–ISPC 2010) concluded that there was much to learn from Challenge Program's application of an experimental design to the evaluation of the relative effectiveness of IAR4D. However, a fully randomized control trial comparing conventional extension and IAR4D would probably have doubled the budgetary requirements and would have been logistically difficult. As such it would have been difficult to justify the costs even if the conventional system could have been specified as counterfactual (CGIAR–ISPC 2010).

Two findings from the evaluation have important implications for methodology and for CGIAR engagement:

1. The Challenge Program experience highlights the difficulty of specifying, ex ante, a reasonable set of accountability targets within an appropriate timeframe, especially in an area like agricultural research and of deciding on a course of action when those targets are not met.
2. The view of the panel is that CG Centers have little future role in IAR4D implementation—that is, the formation of innovation platforms—beyond this piloting or developmental phase.

The panel recommended that two more years for a proper test of the IAR4D concept were required within the context of key revisions of the research plan.

This rigorous experience was largely confined to the Challenge Program because it was trying to prove the concept of IAR4D as an improved path to impact versus other technology adoption programs; none of FARA's other flagship programs, such as the Dissemination of New and Agricultural Technologies in Africa, need such a rigorous research design to evaluate impact.

Lessons for the Design and Management of M&E in Agricultural R&D Institutions

Significant advancements have been made in the development and application of tools and approaches for tracking performance of agricultural R&D systems. The CGIAR Centers managed to institutionalize the practice of evaluation, with the majority of them having fully fledged programs responsible for impact analysis. The adoption of the medium-term plan template as a management tool further elevated the prominence of implementation M&E. Two conclusions can be drawn from the CGIAR experience:

1. The CGIAR Science Council policy helped advance the application of tools and approaches developed from economics and social, biological, and physical science to measure the impact of R&D investment.
2. By adopting a single organizing framework around the medium-term plan, the Science Council promoted the concept of managing for impact or results-based management, a concept hitherto less appreciated in a research setting.

Owing to the paucity of data and complexity of systems, simplifying assumptions sometimes led to overestimation of the impact of R&D investments. Improvements in tools and approaches in the recent past have addressed this problem (Duflo, Glennerster, and Kremer 2006). Alston et al. (2011) have recently provided new estimates of the rates of return to agricultural R&D in the United States that are significantly lower than previous estimates. Lengthened R&D lags, analysis of spillovers and spill-ins by region and size of State, and the relatively low returns to own-research by the U.S. Department of Agriculture could be an example that raises interesting questions for research organization in Africa.

African agricultural R&D institutions (FARA, the SROs, and the NARIs) have been less successful in their endeavor to institutionalize M&E. This situation is in part attributable to the level of planning, and identifying what to monitor and evaluate. As indicated earlier, the analysis by IFPRI informed the restructuring of ASARECA into program areas; however, the individual program strategies did not employ a similar approach to arrive at the priority agricultural research and investment areas. Instead, ASARECA used an inclusive stakeholder consultation process, culminating in a ranking of identified constraints in order to generate the programmatic investment areas. Arguably, a more rigorous analysis based on the framework developed by IFPRI would yield a similar set of priorities. Nevertheless, it is instructive to note that the detailed analysis prioritized investments that enhance productivity of food staples as opposed to the dominant view in the 1990s that public-sector investments should be focused on traditional export commodities (Omamo et al. 2006). Likewise, evidence from the CAADP stocktaking process suggests that the modeling work by ReSAKSS generates a more objective baseline and target values for key indicators at input, output, and outcome levels.

At the operational level, both FARA and the SROs have well integrated results frameworks that facilitate performance reporting for program management. The real challenge to these institutions, however, is that of measuring and reporting impacts. The SROs in particular could borrow a leaf from the reformed CGIAR and adopt a thematic approach to ex post impact assessment. Within the broader Global Forum for Agricultural Research (GFAR) framework, ILAC could backstop the SROs in the design and management of their impact evaluation initiatives. At the country level, the proposed expanded use of ASTI data and sharing of their experience with the NARSs would help address the challenges of data collection and analysis.

The above cases highlight a few lessons:

- the introduction of new M&E activities requires a strong champion at the top of the organization;
- the choice of instruments and supporting information systems requires expertise and a user perspective;
- ex post evaluation needs to be considered in the research design of programs and projects from the start; and
- retention of expertise is a continuing problem for institutions, but expertise is increasingly remaining in Africa.

Consistent application of these tools remains a challenge and, as a consequence, there is no standard approach to assessing the performance of agricultural R&D systems. The framework developed in this paper offers some pointers to the possible way forward. The CGIAR is also challenged to ensure accountability for strategic impact, quality, and relevance of program, managerial, governance, and financial performance and resource mobilization.

8. CONCLUSIONS AND EMERGING CHALLENGES FOR RESEARCH

The preceding discussion has covered both a broad period of history and a wide range of themes. The target audience is the reader who wants the perspective of where we've come from and where this is leading (and where detail has been sacrificed, it is hoped that references to the literature will allow the reader to follow the tracks). Where policymakers have required evidence for decisionmaking—whether technical, policy, or institutional—researchers have responded with evaluation tools: from economic assessment and priority-setting tools to multi-criteria assessment of impact on social and sustainability goals. The tools of adaptive management have helped systems respond to new challenges, often expressed as a need for "reform." Participatory methods for priority setting through evaluation have proven their value in addressing concerns of poor people and in enhancing efficiency.

The final section on the state of tools at the national, subregional, and continental level for Africa is a description of a set of imperfectly interconnected systems and subsystems in an emergent state. In spite of the efforts of donors and African leaders to construct legal institutions, planning frameworks, and structured responsibilities, the "system" is likely to emerge through self-organizing behavior among countries, networks, and local cross-boundary trade. The implications for research evaluation are as follows. First, with subregional research programs coexisting with the networks of CGIAR Centers and new CRPs, planning and evaluation approaches will likely have to become more flexible to deal with nonlinear outcomes and unplanned feedback loops. Second, with organizations and institutions creating new forms of partnership, performance management will have to identify the sources of individual and institutional productivity and study ways to strengthen those sources. Finally, it will become necessary to develop the tools of analysis appropriate to agricultural innovation systems, where accountability is diffuse and the points of intervention difficult to reach.

Bibliography

Acosta, A. S., M. Jones, and R. von Kaufmann. 2005. *The Sub-Saharan Africa Challenge Program: An Experiment in Mainstreaming Institutional Learning and Change.* ILAC Brief No. 9. Rome: Institutional Learning and Change.

Alston, J., G. Norton, and P. G. Pardey. 1995. *Science Under Scarcity: Principles and Practices for Agricultural Research and Priority Setting.* Ithaca, NY: Cornell University Press.

Alston, J., M. Anderson, J. James, and P. G. Pardey. 2011. The Economic Returns to U.S. Public Agricultural Research. *American Journal of Agricultural Economics* 93 (5).

Alston, J. M., P. G. Pardey, J. S. James, and M. A. Anderson. 2009. The Economics of Agricultural R&D. *Annual Review of Resource Economics* 1: 537–565.

Benin, S., M. Johnson, B. Omilola. 2010. *Monitoring and Evaluation (M&E) System for the Comprehensive Africa Agriculture Development Programme (CAADP).* Revised ReSAKSS Working Paper No.6. Washington, DC: International Food Policy Research Institute.

Benin, S., J. Thurlow, X. Diao, H. Kalinda, and T. Kalinda. 2008a. *Agricultural Growth and Investment Options for Poverty Reduction in Zambia*. ReSAKSS Working Paper No. 19. Washington, DC: International Food Policy Research Institute.

Benin, S., J. Thurlow, X. Diao, C. McCool, and F. Simtowe. 2008b. *Agricultural Growth and Investment Options for Poverty Reduction in Malawi.* ReSAKSS Working Paper No. 18. Washington, DC: International Food Policy Research Institute.

Benin, S., Thurlow, X. Diao, A. Kebba, and N. Owfono. 2008c. Agricultural Growth and Investment Options for Poverty Reduction in Uganda. ReSAKSS Working Paper No. 17. Washington, DC: International Food Policy Research Institute.

Breisinger, C., X. Diao, J. Thurlow, and R. M. Al-Hassan. 2008. A*griculture for Development in Ghana: New Opportunities and Challenges.* ReSAKSS Working Paper No. 16. Washington, DC: International Food Policy Research Institute.

Blue, R., C. Clapp-Wincek, and H. Benner. 2009. Beyond Success Stories: Monitoring and Evaluation for Foreign Assistance Results. Independent Monitoring and Evaluation Report. <http://www.usaid.gov/about_usaid/acvfa/060909_beyond_success_stories.pdf>.

Clemens, M. A., and G. Demombynes. 2010. *When does Rigorous Impact Evaluation Make a Difference? The Case of the Millennium Villages.* Center for Global Development Working Paper No. 225. http://www.cgdev.org/files/1424496_file_Clemens_Demombynes_Evaluation_FINAL.pdf

CGD (Center for Global Development). 2006. *When Will We Ever Learn? Improving Lives through Impact Evaluation.* Report of the Evaluation Gap Working Group. <http://www.cgdev.org/files/7973_file_WillWeEverLearn.pdf>.

CGIAR (Consultative Group on International Agricultural Research). 2011a. CGIAR Evaluation Policy (Draft 2). Rome.

_____. 2011b. CGIAR Policy for Independent External Evaluation (Independent Evaluation Arrangement IEA) for consideration of the 6th Meeting of the CGIAR Fund Council Rome November 8–9) <https://www.cgiarfund.org/cgiarfund/sites/cgiarfund.org/files/Documents/PDF/fc6_cgiar_evaluation_policy_22Oct.pdf>.

CGIAR Science Council. 2008a. *Strategic Guidance for Ex Post Impact Assessment of Agricultural Research.* Science Council Brief Number 22. Rome.

_____. 2008b. *Impact Assessment of Policy-Oriented Research in the CGIAR: Evidence and Insights from Case Studies. A study commissioned by the Science Council Standing Panel on Impact Assessment.* Rome.

_____. 2009. *Defining and Refining Good Practice in Ex-Post Impact Assessment: Synthesis Report.* Science Council Secretariat: Rome.

CGIAR–ISPC (CGIAR Independent Science and Partnership Council). 2010. Report of the Second External Review of the Sub-Saharan Africa Challenge Program (SSA-CP). Rome: ISPC Secretariat.

Chapoy, C. 2011. Moving Beyond Evaluation To Learning and Doing. Mind the Gap Forum. <http://www.impactevaluation2011.org>.

Chilonda, P., C. Machethe, and I. Minde. 2007. *Poverty, food security and agricultural trends in Southern Africa.* ReSAKSS Working Paper No.1. Washington, DC: International Food Policy Research Institute.

Chirwa, E. W., I. Kumwenda, C. Jumbe, P. Chilonda, and I. Minde. 2008. *Agricultural Growth and Poverty Reduction in Malawi: Past Performance and Recent Trends.* ReSAKSS Working Paper No. 8. Washington, DC: International Food Policy Research Institute.

de Brauw, A. 2011. "Learning from Evaluations: The HarvestPlus Orange-Fleshed Sweet Potato Project in Mozambique and Uganda." International Food Policy Research Institute Presentation, Washington, DC, June 17, 2011.

Douthwaite, B., S. Alvarez, G. Thiele, and R. Mackay. 2008. *Participatory Impact Pathway Analysis: A Practical Method for Project Planning and Evaluation.* ILAC Brief No. 17. Rome: Institutional Learning and Change.

Diao, X., S. Fan, S. Kanyarukiga, and B. Yu. 2008. *Agricultural Growth and Investment Options for Poverty Reduction in Rwanda.* ReSAKSS Working Paper No. 21. Washington, DC: International Food Policy Research Institute.

Duflo, E. 2011. *Policies, Politics: Can Evidence Play a Role in the Fight against Poverty?* Washington, DC: Center for Global Development.

Duflo, E., R. Glennerster, and M. Kremer. 2006. Using Randomization in Development Economics Research: A Toolkit. <http://www.povertyactionlab.org/sites/default/files/documents/Using%20Randomization%20in%20Development%20Economics.pdf>.

Ekboir, J. 2011. *What Constitutes Rigorous Evidence for Policy Design, Implementation and Evaluation?* Rome: Institutional Learning and Change.

Ekboir, J., and C. Sette. 2011. *Breakthrough Partnerships: ILAC's Contribution to the Change Process in the CGIAR.* Rome: Institutional Learning and Change.

Elliott, H. 2008. "The Evolution of Systems Thinking Towards Agricultural Innovation Systems." Annotated PowerPoint presentation to IFPRI International Conference, "Advancing Agriculture in Developing Countries through Knowledge and Innovation." Addis Ababa, Ethiopia. <http://www.ifpri.org/events/conferences/2008/20080407.asp>.

Evenson, R. E., M. Rosegrant. 2003. The Economic Consequences of Crop Genetic Improvement Programmes. In *Crop Variety Improvement and its Effect on Productivity: The impact of International Agricultural Research*, R. E. Evenson and D. Golin, eds. Wallingford, UK: CABI Publishing.

Fan, S., M. Johnson, A. Saurkar, and T. Makombe. 2009. *Investing in African Agriculture to Halve Poverty by 2015.* ReSAKSS Working Paper No. 25. Washington, DC: International Food Policy Research Institute.

Gertler, P. J., S. Martinez, P. Premand, L. B. Rawlings, C. M. J. Vermeersch. 2011. Impact Evaluation in Practice. World Bank, Washington, DC.

Govereh, J., E. Malawo, T. Lungu, T. Jayne, K. Chinyama, P. Chilonda. 2009. Trends and Spatial Distribution of Public Agricultural Spending in Zambia: Implications for Agricultural Productivity Growth. ReSAKSS Working Paper No. 26. Washington, DC: International Food Policy Research Institute.

Hassan, R. M. 1998. *Maize Technology Development and Transfer: A GIS Application for Research Planning in Kenya.* Wallingford, UK: CABI Publishing.

Hazell, P. B. R. 2008. *An Assessment of the Impact of Agricultural Research in South Asia since the Green Revolution.* Foreword by SPIA Chair. Rome: Science Council Secretariat.

Johnson, M., R. Birner, J. Chamberlin, X. Diao, S. Fan, A.Nin-Pratt, D. Resnick, L. You, and B. Yu. 2008. Regional Strategic Alternatives for Agriculture-led Growth and Poverty Reduction in West Africa. ReSAKSS Working Paper No. 22.

Jones, H. 2009. *The 'Gold Standard' Is Not a Silver Bullet for Evaluation.* ODI Opinion 127. London: Overseas Development Institute.

Kaplan, R. S., and D. P. Norton. 1996. *The Balanced Scorecard: Translating Strategy into Action*. Canbridge: Harvard Business School Press.

Karanja, D. D. 1990. The Rate of Return to Maize Research in Kenya: 1955–88. MSc Thesis. East Lansing: Michigan State University, Department of Agricultural Economics.

KARI (Kenya Agricultural Research Institute). 1991. Kenya's Agricultural Research Priorities to the Year 2000. Nairobi.

Kristjanson, P., N. Lilja, and J. Watts. 2008. *Rethinking Impact: Understanding the Complexity of Poverty and Change.* ILAC Working Paper No. 6, PRGA Program Working Document No. 25, and ILRI Innovation Works Discussion Paper No. 3. Rome; Cali, Nairobi.

Kusek, J. Z., and R. C. Rist. 2004. *Ten Steps to a Results-Based Monitoring and Evaluation System: A Handbook for Development Practitioner.* Washington, DC: World Bank.

List, J. A. 2011. Why Economists Should Conduct Field Experiments and 14 Tips for Pulling One Off. *Journal of Economic Perspectives* 25 (3): 3–16.

Low, J., M. Arimond, N. Osman, B. Cunguara, F. Zano, and D. Tschirley. 2009. A Food-based Approach Introducing Orange-Fleshed Sweet Potatoes Increased Vitamin A Intake and Serum Retinol Concentrations in Young Children in Mozambique. *Journal of Nutrition* 137 (5): 1320.

Maredia, M. K. 2009. *Improving the Proof: Evolution of and Emerging Trends in Impact Assessment Methods and Approaches in Agricultural Development.* IFPRI Discussion Paper No. 929. Washington, DC: International Food Policy Research Institute.

Mayne, J. 2008. *Contribution Analysis: An Approach to Exploring Cause and Effect.* ILAC Brief 16. Rome: Institutional Learning and Change.

Mills, B. F., ed. 1998. *Agricultural Research Priority Setting: Information Investments for the Improved Use of Research Resources*. The Hague: International Service for National Agricultural Research.

Mills, B. F., R. Hassan, and P. Mwangi. 1995. Maize Programme Priorities in Kenya: Measuring Research-Induced Benefits with Multiple, Spatially-Linked Production Zones. KARI Priority-Setting Project, Nairobi (Draft).

Niven, P. R. 2003. *Balanced Scorecard Step-By-Step for Government and Not-For-Profit Agencies.* Hoboken, NJ: John Wiley & Sons.

Njiwa, D., I. Kumwenda, P. Chilonda, F. Olubode-Awosola, and A. Davids. 2008. Monitoring Trends in Public Spending: The Case of Malawi. ReSAKSS Working Paper No. 9. Washington, DC: International Food Policy Research Institute.

Omamo, S. W., X. Diao, S. Wood, J. Chamberlain, L. You, S. Benin, U. Wood-Sichra, and A. Tatwangire. 2006. *Strategic Priorities for Agricultural Development in Eastern and Central Africa.* IFPRI Research Report 150. Washington, DC: International Food Policy Research Institute.

Omilola, B., M. Yade, J. Karugia, and P. Chilonda. 2010. Monitoring and Assessing Targets of the Comprehensive Africa Agriculture Development Programme (CAADP) and the First Millennium Development Goal (MDG) in Africa. ReSAKSS Working Paper No. 31. Washington, DC: International Food Policy Research Institute.

Oruko, L., 2011. *FARA Monitoring and Evaluation Strategy.* Accra, Ghana: Forum for Agricultural Research in Africa.

Patton, M. Q. 2008. *Utilization-Focused Evaluation*. 4th edition. Los Angeles: Sage Publications.

_____. 2010. *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use.* London and New York: Guildford Publications.

Patton, M. Q., and D. Horton. 2009. *Utilization-focused evaluation for agricultural innovation.* ILAC Brief No. 22. Rome: Institutional Learning and Change Initiative.

Ragasa, C. 2010. *Do Organizational Factors Affect Individual Scientists' Productivity*? *A Comparative and Multilevel Analysis of Nigeria and Ghana Agricultural Research Systems.* Washington, DC: International Food Policy Research Institute.

_____. 2011. *Measuring R&D Performance within an Innovation Systems Perspective: An Illustration from the Agricultural Research Systems of Nigeria and Ghana.* Paper presented at the ASTI/IFPRI–FARA Conference, "Agricultural R&D: Investing in Africa's Future: Analyzing Trends, Challenges, and Opportunities," Accra, Ghana, December 5–7, 2011.

Rasheed S. V, A. Hall, and T. S. Vmsidar Reddy. 2011. *Missing the Target: Lessons from Enabling Innovation in South Asia.* UNU-Merit Working Paper 2011-050. Maastricht: United Nations University Maastricht Economic and Social Research institute on Innovation and Technology (UNU-MERIT).

Renkow, M., and D. Byerlee. 2010. The Impacts of CGIAR Research: A Review of Recent Evidence. *Food Policy.*

ReSAKSS (Regional Strategic Analysis and Knowledge Support System). 2010. ReSAKSS Working Paper No. 6. Washington, DC: International Food Policy Research Institute.

SASHA (Sweetpotato Action for Security and Health in Africa). 2009. <http://sweetpotatoknowledge.org/projects-initiatives/sasha>.

Smutylo, T. 2005. *Outcome Mapping: A Method for Tracking Behavioural Changes in Development Programs.* ILAC Brief No. 7. Rome: Institutional Learning and Change.

Spielman, D. J., and R. Birner. 2008. *How Innovative is Your Agriculture? Using Innovation Indicators and Benchmarks to Strengthen National Agricultural Innovation Systems.* ARD Discussion Paper 41. Washington, DC: World Bank.

Spielman, D. J., and D. Kelemework. 2009a. *Measuring Agricultural Innovation System Properties and Performance: Illustrations from Ethiopia and Vietnam.* IFPRI Discussion Paper No. 851. Washington, DC: International Food Policy Research Institute.

_____. 2009b. "Measuring Agricultural Innovation System Properties and Performance: Illustrations from Ethiopia and Vietnam." Paper presented at the International Association of Agricultural Economists Conference, Beijing, August 16–29, 2009.

Thurlow, J. 2008. *Agricultural Growth and Investment Options for Poverty Reduction in Mozambique.* ReSAKSS Working Paper No. 20. Washington, DC: International Food Policy Research Institute.

USAID (U.S. Agency for International Development). 2011. Evaluation: Learning from Experience. USAID Evaluation Policy. <http://www.usaid.gov/evaluation/USAIDEvaluationPolicy.pdf>

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

sustainable solutions for ending hunger and poverty

Supported by the CGIAR

2033 K Street, NW | Washington, DC 20006-1002 USA
Tel: +1.202.862.5600 | Skype: ifprihomeoffice
Fax: +1.202.467.4439 | Email: ifpri@cgiar.org

www.ifpri.org

FORUM FOR AGRICULTURAL RESEARCH IN AFRICA

PMB CT 173 | Accra, Ghana
Tel: +233.302.772823
Fax: +233.302.773676 | Email: info@fara-africa.org

www.fara-africa.org

The Agricultural Science and Technology Indicators (ASTI) initiative compiles, analyzes, and publishes data on levels and trends in agricultural R&D investments, capacities, and institutional arrangements in developing countries. ASTI is managed by the International Food Policy Research Institute (IFPRI) and involves collaborative alliances with many national and regional R&D agencies.

Jointly convened by ASTI/IFPRI and the Forum for Agricultural Research in Africa (FARA), the conference, "Agricultural R&D—Investing in Africa's Future: Analyzing Trends, Challenges, and Opportunities," brought together experts and stakeholders from the region to contribute their expertise for the purpose of distilling new insights and creating synergies to expand the current knowledge base. The themes under focus were (1) why African governments under invest in agricultural R&D; (2) how human resource capacity in agricultural R&D can be developed and sustained; (3) how institutional structures can be aligned and rationalized to support agricultural R&D; and (4) how the effectiveness of agricultural R&D systems can be measured and improved.

The conference was funded by the Bill and Melinda Gates Foundation and FARA.