AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

INTERNATIONAL FOOD
POLICY RESEARCH INSTITUTE
sustainable solutions for ending hunger and poverty

DIRECTOR GENERAL'S OFFICE

December 2005

IMPACT ASSESSMENT Discussion Paper 25

IMPACT ASSESSMENT OF FOOD POLICY RESEARCH: A STOCKTAKING WORKSHOP

Synthesis Report

Jock Anderson, Maria Soledad Bos, and Marc J. Cohen

2033 K Street, NW, Washington, DC 20006-1002 USA • Tel.: +1-202-862-5600 • Fax: +1-202-467-4439 • ifpri@cgiar.org
www.ifpri.org

IFPRI Division Discussion Papers contain preliminary material and research results. They have not been subject to formal external reviews managed by IFPRI's Publications Review Committee, but have been reviewed by at least one internal or external researcher. They are circulated in order to stimulate discussion and critical comment.

Copyright 2006, International Food Policy Research Institute. All rights reserved. Sections of this material may be reproduced for personal and not-for-profit use without the express written permission of but with acknowledgment to IFPRI. To reproduce the material contained herein for profit or commercial use requires express written permission. To obtain permission, contact the Communications Division at ifpri-copyright@cgiar.org.

CONTENTS

Abstract v

1. State of Play (session chaired by Walt Armbruster) 1

2. New and Old IFPRI (session chaired by Rob van den Berg) 23

3. Beyond IFPRI (session chaired by Susan Offutt) 30

4. The Edge (session chaired by Marc J. Cohen) 35

5. Wrap Up (session chaired by Jim Ryan) 43

References 45

Annex 1: Workshop Program 49

Annex 2: Workshop Participants 51

ABSTRACT

This stocktaking workshop* provided an opportunity for several groups with active interest in impact assessment relating to agricultural policy research to share experiences and views about what constitutes good practice in this field. The sponsoring organizations have had a long-standing concern for the relevance and effectiveness of agricultural and food policy research in general and at IFPRI in particular. That concern has been addressed in past meetings and the time seemed right for a further stocktaking. The focus of this 2004 meeting was on impact assessment experience at IFPRI. IFPRI has, since the mid-1990s, carried out a variety of activities aimed at assessing the impact of its policy research, capacity strengthening, and policy communications programs. The workshop brought together practitioners of such impact assessment work, users of such information, as well as researchers whose activities have been the subject of impact assessment.

The cogency and necessity of such impact accounting work was reaffirmed in general terms. There was constructively critical commentary on the merits of particular approaches and instruments, such as narrative recordings and more quantitative methods of attempting to measure effects of research investments. The perennial issue of challenging counterfactuals was necessarily addressed, and the practicality of experimental and quasi-experimental methods considered. The need for consistency of assessment approaches between ex post studies (which have been the bulk of IFPRI's experience to date) and ex ante assessment efforts that represent an increasing share of the assessment portfolio was also discussed. There has long been a commitment to work towards a strong impact-orientation "culture" within IFPRI; the workshop concluded that, while there has been progress in working toward mainstreaming such a culture, there is still far to go, and efforts must continue in this direction.

* Sponsored by the Economic Research Service of the U.S. Department of Agriculture, the Farm Foundation, and the International Food Policy Research Institute (IFPRI), held at IFPRI Headquarters, Washington, DC, on November 10, 2004.

1. STATE OF PLAY

Session chaired by Walt Armbruster, Farm Foundation

AMBITIONS FOR THE WORKSHOP — JOCK ANDERSON

The sponsoring organizations have had a long-standing concern for the relevance and effectiveness of agricultural and food policy research in general and at the International Food Policy Research Institute (IFPRI) in particular. This concern drives this stocktaking, which addresses good practice at IFPRI (mainly) and elsewhere. IFPRI has, since the mid-1990s, carried out a variety of activities aimed at assessing the impact of its policy research, capacity-strengthening, and policy communications programs. These activities, in recent years, have been coordinated by an independent collaborator who serves as Impact Coordinator.

Several sets of themes are addressed in the workshop followed here by respective questions to which answers are sought and tentatively offered at the end of these proceedings. The objectives of impact assessment at IFPRI presently are:

- Showing how IFPRI's policy research, capacity strengthening, and communications programs contribute to the Institute's mission of reducing hunger and malnutrition in developing countries;
- Achieving improved accountability of IFPRI to its investors;
- Enhancing IFPRI's credibility;
- Improving research quality and effectiveness;
- Ensuring continuing relevance; and
- Promoting strategic thinking and fostering a culture of impact awareness in a learning organization.

Q 1: (a) How well have these objectives been achieved? (b) Are these objectives still appropriate? (c) Should new objectives be articulated? (d) Should greater use be made of randomized and quasi-experimental methods?

IFPRI's impact assessment work includes case studies of IFPRI's impact and policy influence at the country level, and studies on specific research themes (e.g., economy-wide modeling). These studies, following an external review process, are issued in an Impact Assessment Discussion Paper series, which also includes work on attribution and other methodological issues, and on lessons learned.

Q 2: Are further such studies justified?

In addition, the Institute has initiatives under way aimed at institutionalizing impact assessment into IFPRI's organizational culture. These include periodic focus group interviews with senior IFPRI research staff to solicit narratives that describe outcomes, influences, and impacts from research and related activities. *Outcomes* are

measures of the use made of the outputs by clients and partners, *policy influence* refers to the degree to which an output of research has influenced or reinforced a policy, and finally, *impacts* are the effects that the policy has had on its beneficiaries, especially poor or food insecure people. IFPRI's impact assessment team developed a database from the first round of interviews that can generate reports by project and research program for future impact assessments and reporting to donors. As appropriate, the information can be de-linked from the researcher who provided it. A paper summarizing the findings from the first round of interviews was prepared, and a second round was launched in August 2004. The database will be updated as new information is available. IFPRI management has encouraged IFPRI research staff to be alert to anecdotal and other examples of emerging or actual influence and impact from IFPRI research. The impact assessment team serves as a clearinghouse for such reports. The narratives are intended to serve as one vehicle for institutionalization of impact awareness.

Q 3: Is the narratives instrument a worthy approach or should it be modified?

While most of the past effort has been on *ex post* impact assessment, *ex ante* impact assessment has now been launched, initially as a component of HarvestPlus, the Consultative Group on International Agricultural Research (CGIAR) Biofortification Challenge Program.

Q 4: Should there be a wider IFPRI effort on ex ante studies?

IFPRI has periodically sponsored conferences and workshops on impact assessment topics, such as that in The Netherlands in November 2001 on "Assessing the impact of policy-oriented social science research," as well as the present workshop.

Q 5: Are further such general workshops justified?

Tentative answers to these five questions are to be found in the "Wrap-up" Section 5 below.

REFLECTIONS ON IMPACT EVALUATION AT IFPRI – JIM RYAN

Four rationales for the conduct of impact assessment at IFPRI remain as valid today as they did when I began to be associated with the endeavor in 1998. They are basically to improve:

- accountability and credibility;
- quality and relevance;
- program and project design and implementation; and
- future planning and prioritizing.

To a significant extent, the primary purpose of impact studies determines the appropriate approach. If accountability is the major reason for evaluation, the evaluator may choose programs or projects purposively, rather than randomly. Choosing the more

"successful" candidates may more convincingly justify the investments in the institution to the public, clients, and donors. However, such "cherry picking" may not be as informative to an institution that is interested mainly in improving its quality, relevance and effectiveness. In such instances sampling "failures" as well as "successes" may offer more insights.

Ryan and Garrett (2003) have distilled lessons from the conduct of a number of impact studies commissioned by IFPRI. Rather, I will try and reflect further on what we might have done differently or might have done that we did not, over the past four years in impact assessment at IFPRI.

Background

IFPRI's Current Strategy and Approach

The early imperative for impact evaluation in the mid-1990s was to enhance accountability to IFPRI's donors in order to justify the wisdom of their investments. The accountability imperative arose because of increasing competition for declining agricultural research and development funding from donors, and policy research had to compete with alternative investments. The hope was that, by providing quantitative estimates of the economic benefits, especially to the poor, one would be able to demonstrate that policy research had comparable impacts to other research themes, as documented by Alston et al. (2000). Ultimately, IFPRI did not pursue the more "global" quantitative approach. Rather, IFPRI undertook a series of case studies, beginning in 1998, and came to adopt a mixture of quantitative and qualitative approaches to assess impact primarily at the project level. IFPRI's 1997 symposium on impact assessment concluded that case studies were the appropriate way to articulate, measure, and document the impact of economic policy research.

The Board of Trustees approved an operational strategy to institutionalize impact evaluation at IFPRI in 2000. The process was discussed in detail with all staff at the IFPRI Internal Program Reviews in both 2001 and 2002. In 2001, IFPRI's management initiated a number of pilot exercises involving *ex ante* impact evaluation on new projects as a component of the new strategy. IFPRI also began to go beyond the project-level to conduct evaluations of some its thematic research programs. The first such study is the examination by Alwang and Puhazhendhi (2002) of the impact of IFPRI's multiyear, multicountry research program on microfinance. The second was by Ryan (2003) on agricultural projection modeling at IFPRI. The third was by Anderson (2003) on economy-wide modeling.

The case studies provided *ex post* evaluations somewhat distant from the daily operations of staff. IFPRI seeks now to incorporate some aspects of impact evaluation in all its research activities to improve IFPRI's ability to operate as a learning organization. The narratives work described later in this workshop is one important instrument in this work.

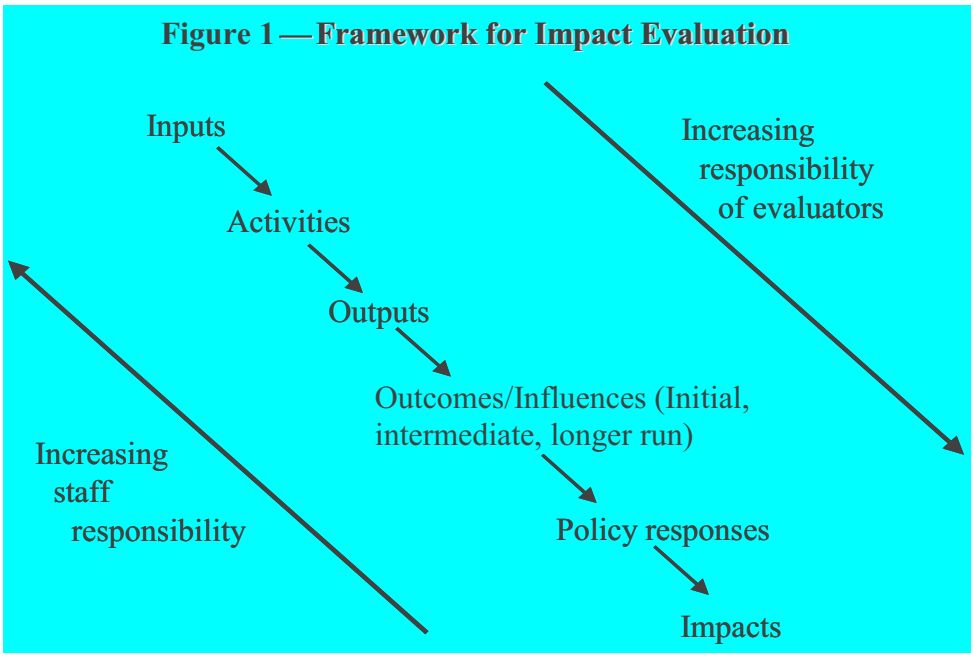
IFPRI's current approach to impact evaluation is to categorize the products from economic policy research and related activities as outputs, outcomes/influences, policy responses and welfare impacts (Table 1). *Outputs* are activities or effort that can be expressed quantitatively or qualitatively. *Outcomes* or *influences* are measures of the use that clients or partners make of the outputs. They reflect the value placed on them as intermediate products, which in turn provide inputs into the policymaking process. Outcomes and influences can be usefully separated into initial, intermediate, and longer-term. *Policy responses* imply a degree of attribution of the effects of the intermediate outputs and outcomes/influences on the formulation of new or reinforcement of existing policy. *Impacts* are measurable effects of the attributed policy responses on the well-being of the ultimate beneficiaries of the research, namely the poor, the food and nutrition insecure, and the environment. It could also include perceptions of peers and policymakers about such impacts.

Table 1. Some Indicators of the Products of Policy Research

Outputs	Outcomes/Influences	Policy Responses	Impacts
Publications • number and type • refereed/non-refereed *Methodologies* • description • value-added *Training* • number of trainees • extent of training • duration of training • number and type of manuals *Seminars/Symposia/ Conferences* • number • type • number of participants *Press Releases* • number • type *Press Conferences* • number • type *Capacity-Strengthening of Partner Institutions*	*Publications* • citations, use in curricula, circulation numbers, sales, requests, web hits *Methodologies* • use of new methodologies *Training* • trainee promotions • number of others trained by IFPRI trainees *Seminars/Symposia/Conferences* • number of policy-makers present and influence on policy • invitations to IFPRI staff to present keynote and other papers at other meetings—number, organizations, and whether expenses are paid *Press Releases* • number of press releases published and in what fora; letters to editors, news articles, and editorials spawned as a result *Press Conferences* • number of press articles that resulted and in what fora *Capacity-Strengthening* • Invitations to IFPRI staff and management to be on committees adjudicating policy changes in partner organizations and countries • Refereeing assignments of IFPRI staff,	Changes in policies attributable to policy research Reinforcement of existing policies Implementation of policy changes Changes in institutions	Reduced poverty Improved food and nutrition security Sustained livelihoods of the poor Enhanced natural environment

	requests for additional research in response to earlier outputs • Degree of success in acquiring additional resources for policy research at partner institutions		

These products are generally sequential. Evidence becomes more difficult to assemble as one moves from outputs to impacts. Generally, the responsibility of staff and management for documentation and evaluation decreases on the same continuum, whereas the role of independent peer evaluators increases (Figure 1).[2]



For this framework to function, staff must record output and outcome/influence indicators as a matter of course. Increasingly, this is happening at IFPRI. These indicators will relate directly to milestones and achievements in workplans at the beginning of the year, and on the rolling, three-year Medium-Term Plans. IFPRI staff will also record policy responses, subject to subsequent verification by independent peer impact evaluators, which may be captured effectively in narratives. This is not meant to be a top-down compliance approach, but instead a performance management approach. Eventually it could become a regular part of staff evaluation, although presently it is not.

Investors in public research and development are no longer satisfied with activity- and output-based progress reports. They expect outcome/influence and impact evaluation. That is, objective assessments of the actual effects of the funded program on the target population (Easterling 2000). For research institutions to deliver on this requires responsibility and accountability at the staff level. Suitable databases of indicators of outputs, outcomes/influences, and policy responses need to be developed

[2] The graphic is not meant to imply the policy process is linear but only to portray the responsibilities of staff and evaluators in assessing impact.

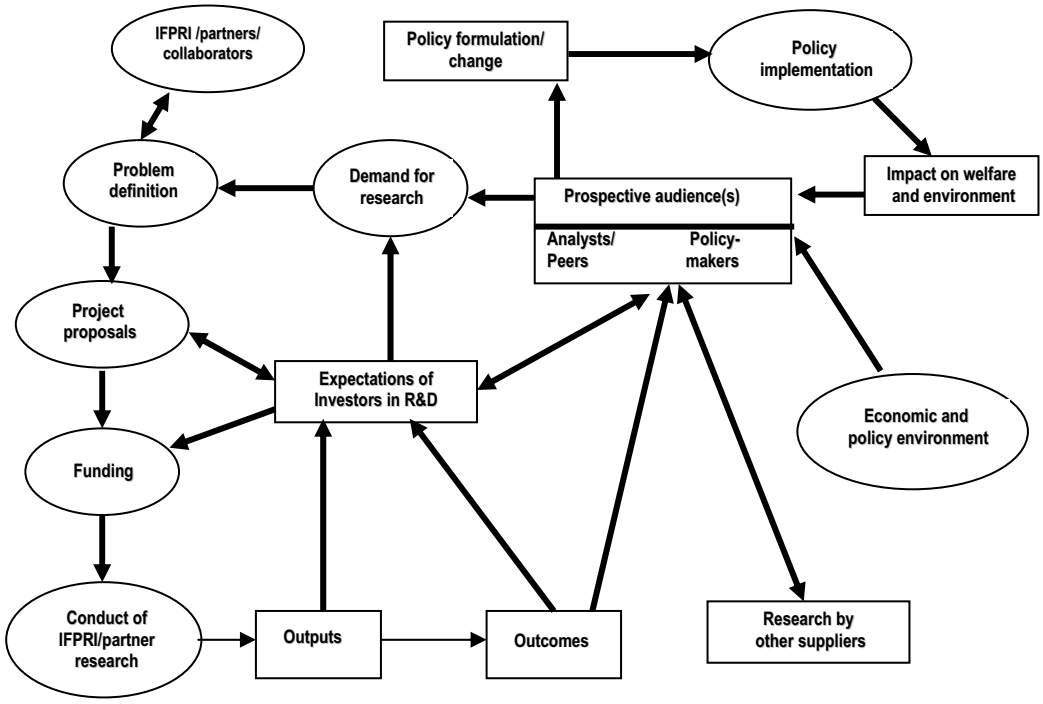
and maintained so they become a sustainable part of the corporate memory that is not lost when individual staff members leave. Given the often long period between the conduct of economic policy research and the generation of real impact, such databases are imperative.

Indicators at staff level are then aggregated to the appropriate project-, program- or institute- levels. Benchmarks are important for all of these. These can be before-and-after comparisons that document the gap between goals or milestones and actual achievements. A framework that ties outputs to processes and assumptions can provide a useful way of linking such *ex ante* with *ex post* impact assessments. This is now a feature in the CGIAR (Balzer and Nagel 2001). For a more comprehensive impact assessment comparisons with best practices of others (i.e., benchmarks with-and-without IFPRI) are desirable.

The Policy Process

As Sutton (1999, p.32) indicates, it is important to recognize that policymaking is a political process and not, as many economists might prefer to describe it, an analytical or problem-solving one. Hence, effective impact evaluation requires an understanding of the policy processes. Figure 2 is a schematic representation of the policy process and the points at which impact evaluation of economic policy research would primarily focus. Of course, each country differs in the detail and the figure is just a generic representation for purposes of exposition.

Figure 2. The Policy Research Process and Evaluation Nodes[a]



[a] The rectangles represent nodes where evaluation efforts would primarily focus. The ovals are the actions or influences.

Policy formulation and change is subject to a complex array of causes and influences. These include the domestic and international economic and policy environments, including importantly the voting public, civil society, and other interest groups. Economic research by institutions such as IFPRI and its partners is also expected to have influence, as is that of other research suppliers. Investors in research such as national governments, aid agencies, and international institutions also have a stake in and an influence on the policy process.

Sutton's (1999) contention is that the linear model of policymaking characterized by objective analysis of options and separation of policy from implementation is inadequate. Instead, it is argued to be "a chaos of purposes and accidents". Policymaking is interactive and nonlinear. Separating policymaking from implementation of policy is not appropriate when trying to assess impact, as without effective implementation there is unlikely to be impact. In separating the two there is a danger that the former is viewed as the realm of politicians and the latter the task of bureaucracies. In reality continuous feedback and feedforward is a feature of the process. It is interactive and highly nonlinear, involving a diversity of actors (Scott 2000).

This complexity makes impact evaluation of economic policy research a somewhat heroic enterprise, especially when one is mainly concerned with accountability of and attribution to individual institutions. In such cases one usually begins from the "supply-side" at the conduct of the research by the institution, its outputs, outcomes and influences on policy formulation and change via various audiences of policymakers, policy analysts who serve them and professional peers, taking due account of the other factors that affect these also (see Figure 2). Among the factors are the predisposing economic and policy environments and the research conducted by other suppliers, both historically and contemporaneously. The latter is especially relevant in the context of being able to attribute outcomes and policy changes to particular institutions. Then the economic welfare and environmental impacts of the research being evaluated are assessed. Of course, effective policy implementation is a necessary step in achieving impact. It is moot whether this phase of the process has ever been or should be included as a component of impact evaluation. It is obviously a key part of enhancing impact. We will return to this later.

If attribution of impact is not a primary reason for undertaking assessment, then one has better scope to explore the issue from a "demand-side" perspective. In this instance one would commence from a particular policy formulation/change event or related events and work backwards (see Figure 2) to the various research institutions that conducted research on the policies concerned. By focusing particularly on the audiences for the research information and eliciting from them what were the major influences on decisions, one can establish the role of the various research suppliers. This is a more satisfying and perhaps objective way to assess impact and an approach favored by the participants in at least one focused workshop (TAC Secretariat 2001, p. 19). However, it may not be as effective in drawing out lessons for individual institutions in enhancing future impact as a more "supply-side" approach.

Babu and Mthindi (1995) bemoan the fact that decision makers often do not use seemingly relevant information in making policy decisions. They point out that too often decision makers are involved in "firefighting," with little time for informed decisions. Paucity of data is also often a handicap to policy formulation. Babu and Mthindi separate the benefits of policy research into pre- and post-decisionmaking benefits. The former involves improved processes related to capacity building and institutional strengthening. The latter are evaluations of the primary and secondary impacts of the policies that emerge. Babu and Mthindi measure costs and process benefits, but do not attempt to measure impacts.

Garrett and Islam (1998) suggest that social science research evaluation should only look at outputs, processes, and potential outcomes, rather than focusing on actual policy outcomes. They maintain it is difficult to establish a direct link to the policy impact of social science research and that often the research contributes to a body of knowledge that policymakers access when and if they see fit. According to Garrett and Islam, there are four features that policymakers find useful: (1) research quality, (2) conformity to expectations, (3) action orientation, and (4) challenge to the status quo.

It is contended here that evaluating the quality of the research output and the processes by which a research institute carries out and communicates its research findings is a necessary but not a sufficient condition for judging impact. Garrett and Islam maintain that it is sufficient. One must also look at post-decision impacts if an institution is going to be able to differentiate its product from others and sustain funding support in this era of increased accountability and contestability.

Some maintain that studying dissemination measures is a good proxy for impact. The disembodied knowledge that arises from economic policy research has the characteristics of a public good. It is non-rivalrous and once public, non-excludable (Stiglitz 1999). Stiglitz differentiates between tacit knowledge and codified knowledge. The former involves horizontal knowledge transfers using special methods such as apprenticeships, secondments, study tours, learning by doing and institutional twinning arrangements. Codified knowledge on the other hand involves vertical transfer methods using central repositories such as libraries and electronic means that are accessed as required by interested parties.

Some Lessons and Hardy Perennials

Based on my involvement with IFPRI in impact assessment there are a number of unresolved issues and puzzles that require further attention. In discussing these I will try and distill the key elements and how I think we might proceed in future.

Institutional Learning and Impact Assessment

We tried to draw lessons from the various impact assessments that were conducted. The lessons were of two types: (1) How to improve the conduct of impact

studies from a methodological perspective; and (2) How to enhance future impact. Ryan and Garrett (2003) have summarized the lessons learned to date.

There is a new imperative in the CGIAR related to the concept of institutional learning and change (ILAC). It contends that per se, traditional economic impact assessment (EIA) is not as valuable to institutions, investors and stakeholders as ILAC approaches and that this explains why EIA has not been as effective as it might have been in influencing internal and external research management and decisionmaking (e.g., Springer-Heinze et al. 2002). Some even go so far as to say there is a tension between the two approaches, which surfaced at the 2002 Costa Rica impact assessment conference of the International Maize and Wheat Improvement Center (CIMMYT) and the Standing Panel on Impact Assessment (SPIA) of the CGIAR Science Council (Horton and Mackay 2002; Mackay and Horton 2003). More attention is required to impact chains and pathway analysis to better understand the processes that are portrayed in Figure 2 above if impact-oriented institutions are to be cultivated. EIA alone will not achieve this.

With the integration of the erstwhile International Service for National Agricultural Research (ISNAR) programs dealing with institutional innovation and change in agricultural R&D systems into IFPRI, there is an excellent opportunity for IFPRI to develop leadership in this field. This would allow IFPRI to complement its concerns about improving the links between its research and the formulation of policy, and the equally important issue of improving the links between policy formulation and its implementation.

I contend that, at IFPRI, we have managed to make the accountability aspects of EIA complementary to the ILAC aspects. The author's presentations to the IFPRI Board, his paper at the Costa Rica conference, the paper by Ryan and Garrett (2003) and the Scheveningen impact assessment workshop in 2001 (Ryan 2002) illustrate the latter, and the various case studies either commissioned or conducted by him illustrate the former. Proponents of the ILAC approach tend to regard the two as competitive but this need not be the case. Also, ILAC uses somewhat new concepts and a new jargon associated with them that seems to imply that unless those conducting evaluations use them somehow there will not be legitimacy to any ILAC implications that are drawn from them. Whilst arguably the IFPRI Board and management might not have accepted full ownership of the lessons that were drawn from the various EIA case studies, these did represent an attempt to go beyond traditional EIA.

ILAC approaches are intended primarily to improve internal institutional efficiency and performance management. EIA is primarily intended to improve accountability to external investors and stakeholders. Both aspects deserve more attention and it is good to see that SPIA has recently conducted a survey of donors to elicit their expectations of impact assessments in the CGIAR. The intention is to see if impact assessment (IA) can better respond to donor demands in future, including their views on EIA and ILAC aspects.

Institutionalization of Impact Evaluation

IFPRI has improved its impact orientation over the past few years. However, there is a long way to go before one could claim that IA has been "institutionalized". For example, little has been done to more explicitly link the IAs with the assessment of future institute priorities, one of the four hoped for rationales for conducting such studies. An IA culture is growing but slowly. I submit that until project/program *ex ante/ex post* IAs are more explicitly linked to staff workplans and evaluations with a buy-in by the Board, management, and staff, neither objective will be achieved. It is timely to note in this context that the CGIAR Science Council has proposed that CGIAR center Medium-Term Plans and logframes in future be more explicit about outputs, outcomes, impacts, and associated milestones so that more effective monitoring and evaluation will be possible. Hence, I believe IFPRI would be well advised to move forward on linking program/project IA with staff workplans and performance management. The retrospective narratives from staff are a step in the right direction for this to occur, but they will require introduction of explicit and regular peer scrutiny and validation/corroboration before they can be regarded as credible. Additionally, provision for annual updates has to be addressed. At present, they have been somewhat *ad hoc.*

A paper on all of this was drafted at the request of IFPRI's Senior Management Team and presented to them in February 2002. However, there was weak follow up. The Working Group on Impact Evaluation (WGIE) at IFPRI discussed the issue of the institutionalization of IA and staff evaluations at its February 2002 meeting but again there was little follow-up. As long as IA is seen by most as a necessary evil, which is best managed by an outside consultant, I am afraid that its institutionalization will remain elusive.

Focus on Final Impacts versus Outputs, Outcomes, and Influences

I understand that the CGIAR centers now have regular access to the various scientific citation indices from the Web of Science. This will enable regular and explicit documentation of an important component of the outcomes from outputs than was the case previously. In the conduct of recent IAs we tried to access this but could only do so for one or two of them. Until we are able to have updated bibliometric citation benchmarks from peer institutions and individuals, as Pardey and Christian (2002) compiled, against which to compare contemporary outcomes, it will not be possible to effectively monitor this crucial ingredient for ultimate impact. The new Web of Science database access should enable this to occur but will require IFPRI to assemble its own capability to download and analyze information as a normal part of monitoring and evaluation (M&E) processes both at the individual staff, project, program and divisional levels.

The Natural Resources Institute of the UK has proposed a "Balanced Scorecard" approach to impact assessment of research institutions, where the goals of each of the structural elements of the institutions are established initially, and performance indicators devised and subsequently evaluated (Smith and Sutherland 2002). These goals and

indicators are generally different for each of the institutional elements such as finance, staff, clients, and the internal business. Whilst such approaches are valuable as part of institutional M&E oversight, it is not clear to me that they can be regarded as alternatives to EIA in the quest for improved accountability and impact, especially in terms of the ultimate clientele for policy research, namely the poor, the food insecure, and the environment. They obviously have value in improving internal efficiencies and performance of the institution but are arguably less convincing in establishing the credibility of external performance. Guidance from IFPRI's investors would be helpful here.

Project versus Program or Thematic Evaluations

Thematic impact assessments of whole bodies of work such as the recent ones by Anderson (2003) and Alwang and Puhazhendhi (2002), allow the analysts to reflect more on the international public good aspects of IFPRI's outputs. Some argue that this should remain at the heart of IFPRI's focus and hence that IA should primarily concentrate at the thematic or program levels rather than at the project or country level such as the studies by Ryan (1999a, b). The latter research is more in the nature of private goods or perhaps at best national public goods, best left to other suppliers to undertake in the first place.

Clearly project- or country-level IAs allow the assessor a better prospect of eliciting from stakeholders and clients their perceptions of the value and influence of IFPRI policy research in terms of subsequent actual policy changes. To the contrary, IAs of long-term research on themes by their nature have to rely more on citations analyses and derived demands for access to outputs such as web downloads and publication requests.[3] It is generally more difficult to track influence on specific policy changes at the thematic levels and hence less likely that EIA of final impacts will be possible at this level. There is hence a tradeoff here, and there is no easy answer as to what the appropriate balance should be. I would certainly tend to favor keeping a balanced portfolio of IAs at the various levels rather than focusing on one or the other.[4]

A related issue is to what extent investors would be more comfortable with IFPRI moving away from country technical assistance type projects and more towards cross-cutting thematic research closer to the international public goods end of the spectrum. This has been and remains a hardy perennial issue at IFPRI. If IFPRI moves more upstream in this manner, I believe it then will be more difficult to both document and measure final impacts. Then investors will have to be satisfied with intermediate

[3] IFPRI is to be applauded for making available on its website in recent years social accounting matrices, computable general equilibrium models, and household survey datasets to bona fide researchers. These are important international public goods. The increased attention to IA may have played a role in this.

[4] In his assessment of the impact of agricultural projection modelling, Ryan (2003, p. 38) found that IFPRI publications using the IMPACT framework were more influential on respondents to the survey he conducted than professional journal articles. This raises a question as to whether there may be a tradeoff in the pursuit of institutional impact versus the professional growth interests of staff members of IFPRI. This deserves more attention in future.

products, which capture outcomes and influences rather than attributable policy responses and impact.

To quote from a recent survey of policymakers by Dinello and Squire (2002, p. 17) under the auspices of the Global Development Network (GDN):

"Several general results emerged clearly from the survey:

- *To a perhaps surprisingly high degree, policymakers in this survey acknowledged their tendency to draw on research results, attribute importance to local research and the research community when formulating policy, and be satisfied with the research-policy link.*
- *The policymakers also tended to value research that is characterized by policy relevance, timeliness, and high quality.*
- *The policymakers attributed high value to focused specialized research that is well tuned to their interests, rather than general research, regardless of who produces it (government or independent research institute).*

These preliminary findings are basically favorable. According to them, research-based policymaking happens more often than might have been expected. From a researcher's perspective, access to policymakers depends on their ability to produce policy-relevant and high-quality research and attract attention to its results. The policymakers' high rating of "specially commissioned studies" as a source of information confirms their preference for sharply focused policy-relevant research. According to the survey's preliminary findings, enhancing the policy relevance and quality of the research is the best route to bridging research and policy.

These findings are consistent with the lessons drawn by Ryan and Garrett (2003). They also reinforce the likelihood of the tradeoffs mentioned previously from a focus on specific or country-level versus thematic research in terms of the likelihood of subsequent attributable policy responses and influence.

I agree with Anderson (2003, p. 39) and Ryan (2003, p. 39) that more responsibility should be placed on IFPRI staff to compile databases of citations of their work in the "grey literature" such as in university syllabi, academic papers, Ministerial speeches, etc. Independent assessors cannot be expected to initiate this process at the time IAs are commissioned. Again, this is related to the issue of institutionalizing impact assessment at IFPRI and relating it to staff work-planning and performance assessment.

Operational Strategy

This was discussed at IFPRI's annual internal program review in 2001 and 2002 after the Board of IFPRI approved it in 2000. Is there a need to revise the strategy in the light of subsequent experience? What should be the next candidates for IAs and what new ones should be added to the list compiled for inclusion in the operational strategy? Some issues raised in this paper require consideration within the context of a revised operational plan.

Interactions with Others

There has been interaction with peers, stakeholders, investors, and clients in the definition and conduct of IFPRI's IA work. However, I believe this could be strengthened in future on an *ad hoc* basis. I do not think there is sufficient interest in a consortium of the type proposed originally in 1997 and reaffirmed at the 2001 workshop in Scheveningen. We tried in 2002 but interest was limited. Groups that ought to be included in a wider new net include the Development Assistance Committee of the Organisation for Economic Co-operation and Development, the GDN, the Operations Evaluation Department of the World Bank, and the Overseas Development Institute, among others.

Elicitation and Narratives

More attention is required to survey and interview methods for eliciting the views and perceptions of peers, stakeholders, and clients of IFPRI's research. I felt inadequate in this area and it probably requires some specialist inputs from non-economists. Initial sampling frames for interviews are biased, because with supply-driven IAs the assessor is primarily reliant on the researchers involved in the work being evaluated to compile a list. Of course, one asks the initial interviewees for names of others who might know of the research, but even there one misses those who should have but did not become exposed to the research in the first place.

Should the retrospective impact narratives be continued? If so, should they be done individually in future rather than in focus groups as was the case in the first round? Should there be a questionnaire approach or personal interviews? How do we obtain corroboration in order to ensure credibility? To continue this approach will require substantially more resources than has been the case in the past and I would only favor investing more if it is made an integral part of ongoing staff work planning and performance appraisal.

The Conduct of the IAs

A related issue is the amount of time that IFPRI has allowed for the conduct of external IA case studies. In general, they have been restricted to 30 days. My experience is that this is insufficient if one wants the assessors to do an adequate job of interviewing people in the field in developing countries and writing up a professional paper for the IADP series. This was recently reinforced by Kym Anderson in his IADP (Anderson 2003, p. 39). The process is open ended and time consuming, especially if face-to-face interviews are to be the norm, rather than using mailed questionnaires. I would suggest 40 days is a minimum. In general, I would not favor jointly conducted IAs. We tried it with mixed success and as they were much more expensive than using one assessor, the perceived benefit-cost ratio was much less.

I think that the experience with using non-economists to do the IAs was good, although we only did this once, in the case of the gender and intrahousehold dynamics

IA. At the time of writing the IADP is yet to be published but I found that the author of this case study has raised some important strategic issues for IFPRI in the draft report, as well as providing interesting perspectives about how a non-economist assesses impacts. [Editors' note: this work has since appeared as Impact Assessment Discussion Paper 23 by Jackson (2005).]

The Attribution Hardy Perennial

The more emphasis on accountability as the primary rationale for the conduct of IAs, the greater is the need for attribution of outcomes, influences, policy responses and final impacts to projects, programs, or institutions. Additionally, if institutions such as IFPRI and their investors initiate such studies, then the likelihood is that supply-driven approaches will be the norm rather than demand-driven ones. These involve a potential moral hazard on the part of assessors who are commissioned by the institution being evaluated. It is embarrassing to have to lead the interviewees in questioning to ensure that IFPRI's role is at least discussed. There is a danger that in this process the roles of alternative suppliers of policy research and advice will be underplayed.

I would like to suggest that IFPRI commission one or two demand-driven approaches where the starting point is a major policy change. Assessors would be asked to work backwards from these to distill the various influences on the decisions, including the contributions of the various research institutions involved, including of course IFPRI (Figure 2). There would be less emphasis on attribution to individual institutions and more on narrating the disparate contributions of all the players. Some of the case studies commissioned in the Bridging Research and Policy project of the GDN are of this type. Others are contained in the excellent synthesis of the contribution of human nutrition research and advocacy to the formulation of nutrition policies by Gillespie et al. (2001) and of the role of biological and economic policy research in the development of policies on tuna fishing in Australian waters by Campbell and Squires (1998). These demand-side approaches rely heavily on retrospective narratives in the case studies and IFPRI could learn a lot from these. Whether investors will be satisfied with them in terms of the accountability of individual institutions is moot though. IFPRI needs to poll its investors to find out.

Conclusions

In summary, the case studies instituted by IFPRI in the past five years have proved to be effective in articulating impact, thus satisfying the accountability imperative, the primary rationale. The case studies in this series have highlighted lessons that the institute can use in the *ex ante* planning and conduct of its future research and related activities. This was the second of the four rationales for conducting impact evaluation.

Progress also has been made in the past five years from the conduct of various case studies of the impact of economic policy research in drawing lessons for the future. Some progress has also been made in the development of methods for quantifying impact

in economic terms. However, a number of issues remain. These include attribution, measurement and the enhancement of impact.

"Demand-side" approaches seem preferable to "supply-side" ones. The former uses major policy events as the starting point and works retrospectively to establish the separate influences of the many research suppliers and other factors on policy responses. It may be easiest, and most logical, to measure joint impacts of various players rather than separating out the contributions of individual institutions such as IFPRI. Impact assessment will also most likely use a mixture of both qualitative and quantitative methods. Retrospective narratives are an essential component of the former, and indeed provide the basis for quantitative estimates, and can help address the elusive issue of attribution. Most importantly, if impact evaluation itself is to be effective in enhancing the impact of research and researchers, the process must be institutionalized. Staff and management should take responsibility for recording outputs, outcomes/influences, and policy responses related to their research. Independent evaluators can verify these and endeavor to translate them into meaningful measures of their impact on economic welfare, and assess what share might be attributed to policy research institutions and their partners. Researchers must see such a system as integral to learning and improving their own actions.

Researchers also have a responsibility to ensure the public dissemination of their findings. To the extent that the independence and credibility of the researcher and the institution are not compromised, a degree of advocacy is also appropriate. With the increased availability of information technology and impetus for participatory democracy and good governance in developing countries, there is increased scope for credible policy research to be accessed by disparate groups and thereby generate public debate and so better inform the policy process. In this context, credible research on the distributional consequences of alternative policies will arguably have more influence and impact than will measures of the implications for economic efficiency.

Finally, we need to continue to undertake more case studies in order to further refine approaches and methods for impact evaluation and help to define "best practices". There is also scope for more multidisciplinary research into policy processes in order to better position policy research to have strategic influence, as is being proposed in the Food Consumption and Nutrition Division of IFPRI (Haddad and Pelletier 2003). But we should go beyond into bridging the gap between policymaking and implementation. In other words, "bridging policy and action" should complement "bridging research and policy".

COMMENTS BY JAMES GARRETT

In social science we are pushed beyond quantitative measures and beyond just outputs. Measuring impact requires qualitative analysis. A good example is the study of Viet Nam's rice policy change and the contribution of policy research by Ryan (1999). In this analysis the author looks at the "potentiality" of that method to decrease poverty

and it also includes narratives from the main partners and stakeholders on the "perceived" influence of IFPRI.

We have to move towards a social science framework when analyzing impact assessment. In this case, the logical framework of M&E can be very useful.

Policy responses and impact of research are based on many assumptions and affected by many stakeholders. The question remains whether we can hold a specific institution responsible for these.

Also, we should consider if the final impact is the ultimate barometer to measure success. The indicators of the products of policy research (outputs, outcomes/influence, policy responses and impacts) that Ryan develops are very useful in capturing the many components where we can measure successful research. At present, inside IFPRI there is already an internal discussion among staff about the methodology necessary to measure these indicators and that is a good step forward. Furthermore, at some point, investors will have to be satisfied with these intermediate indicators.

We should make sure that donors are also responsible for achieving impact.

Also, worthy of further discussion in achieving "impact" is a discussion of the use of knowledge in the policymaking process. There seems to be a diminution of the use of knowledge by policymakers. How does IFPRI work in a world where policymakers function more on ideology than "facts?" Does it connect with them or with bureaucrats? Except in a few cases, IFPRI work may not dramatically shift paradigms, or provoke large shifts in voter sentiment (in democracies). What are the levers of influence to get others to use our findings then? And how do we legitimate ourselves as the arbiter of knowledge and participant in the policy process?

There is debate at present among IFPRI's researchers as to what is their role in disseminating the results of their research. Some researchers believe that their role is just to present options through their research and not to go any further to achieve policy influence. They adamantly believe that IFPRI should not actively promote one option or the other, and that the role of the researcher is to be an "unbiased" observer presenting alternatives. But some others believe that research per se will not have policy influence and that their role, and IFPRI's role, is to become more involved in disseminating the results. The question is whether all researchers should be involved in disseminating the results or whether there can be a division of labor, in which some researchers will be more involved in disseminating the results and others will be merely researching.

The approach of assessing impact at IFPRI through case studies has evolved from a country focus to public-good/thematic criteria. Still, there has not been any "demand driven" impact assessment study.

Further, we need to answer some key questions: Who should our audience be? Who should we be reaching? What is our role in the policy process?

We should also make sure that the information produced through the impact assessment activities here at IFPRI is used. And this information can be used at three different levels: (1) Internal learning, (2) Other institutions' learning, and (3) To contribute to social science.

It would be particularly important to produce some political and institutional analysis about why IFPRI has not followed up on Jim Ryan's previous recommendations. Basically Ryan recommends in his presentation the same things he did a few years ago. Why has not that happened? This parallels the idea of "provision of information does not always lead to action." Do we hold Ryan, the provider of knowledge, accountable for the fact that IFPRI did not follow up on his recommendations?

CHALLENGES TO ASSESSING RESEARCH AND EXTENSION IMPACTS — WILLIAM M. RIVERA

International organizations in their plans and efforts to assess program impact are confronted with four challenges.

***Models and norms* to be used in assessing Research and Extension (R&E) impacts**

Attribution and linkages in the causal hypothesis are difficult for any impact evaluation. For instance, links from research to adoption to policy recommendations are difficult. Also, links from capacity building to improved research and/or policy formulation are difficult. Certainly, if impact analysis is to be used, certain basic activities must be put into place. Designing an impact system in the program must be assured from the start, followed by periodic monitoring of performance and the final evaluation of results.[5]

Some researchers, e.g. Javier Ekboir of IFPRI, argue that impact analysis should not be used for research evaluation since many factors influence outputs and adoption. Rather research impacts should be analyzed as part of a complex adaptive system that depends on external forces (e.g., markets), the direct and indirect interactions among agents (e.g., researchers, input suppliers, and farmers), and the technology's nature and evolution.

The decision to undertake program impact assessment is thus put in question. Nonetheless, if impact evaluation is pursued, program hierarchies are helpful in clarifying the steps in the program's development (Figure 3). This program hierarchy differs

[5] Monitoring is defined as the process of observing and gathering data to ascertain the quality and efficiency of performance and doing things right. Evaluation is assessing effectiveness, efficiency and sustainability and is also useful to ascertain if projects are doing the right things towards the desired impacts.

somewhat from that in the IFPRI documents and their sequence of *inputs-activities-outputs-outcomes-and-impacts*.

The hierarchy in Figure 3 underlines involvement and the possibility of non-involvement (or reaction) to the project. Also, while it includes material "outputs" (i.e., events or physical products), it emphasizes behavior change in KASA—knowledge, attitudes, skills and aspirations (coined by Claude Bennett, USDA retired). But this behavioral change is not enough; the new KASA must be put, or allowed to be put, into practice. For example, farmers may be trained, and learn to use, certain agricultural supplies that then prove unavailable.

Figure 3. Program Hierarchy

IMPACT
End Results (Outcome)
Practice Change

Type of Change (Outputs/KASA)
Reaction
Involvement
Activities
Inputs

Furthermore, the new paradigm—of privatized/decentralized/pluralistic planning and decisionmaking—complicates research and activity planning, as well as the uptake of policy recommendations resulting from programs.

Recognize and clarify the diverse <u>types of impact</u> that contemporary development programs are presumably set up to seek

Impacts are diverse and are valued differently by different stakeholders. In practice, measuring capacity impacts is difficult. Impacts will depend on the characteristics and goals of the program, the quality of performance, the output of material products, and/or behavioral learning. If behavioral learning is put into practice, then there is hope for an impact. These impacts might include:

- Policy impact = policy becomes revised or changed
- Economic impact = cost/benefit ratio or income generation improved
- Social impact = empowerment, equity, poverty alleviation, or community development are promoted
- Organizational impact = improve performance/finances, linkages, institutional sustainability
- Natural resource management impact = management of natural resources and environmental sustainability are improved
- Human resource development impact = human capital advanced through education and training

How contemporary trends are affecting the development of R&E programs and their impact

- Globalized economies are changing the way we look at impact. Countries are forced to examine more carefully their comparative advantage and the global marketplace.
- Also, whereas public institutions operating in a top-down manner were the paradigm for past development efforts and R&E investments, new pluralistic systems of R&E innovations are mandated by a new set of strategies — decentralization, demand-led R&E, and privatization — which also affect the types of impact sought.
- Demand-driven program goals require new thinking and actions to promote end-user involvement in development.
- Outsourcing and privatization introduce alternative dimensions to program development.[6]

Review the varied conceptual frameworks and decide which of these conceptual orientations should guide impact assessment

There is no clear agreement on the scope of the institutional framework for impact assessment. A plethora of concepts and nomenclatures currently compete for allegiance. It is not my purpose to take sides on this issue, but rather to highlight what appears to be an imperative to confront these overlapping but distinctive concepts of R&E development. I suspect that given the terms — National Agricultural Research Systems (NARS), National Systems of Innovation (NSI), Agricultural Innovation Systems (AIS), and Agricultural Knowledge and Information Systems for Rural Development (AKIS/RD) — that each of us here will have different meanings and distinct views as to their advancement. This is a serious challenge as it tends to divide rather than unite.

In general, NARS and NSI refer to national systems of research and extension. AIS reflects a more integrated agricultural research/extension orientation, recognizing other support and relevant systems. AKIS/RD also seeks to promote an integrated approach involving agricultural education, research and extension, as well as other agricultural development support systems. And to muddy the waters, more recently development professionals have begun to highlight a new challenge: to diffuse development information on non-agricultural rural needs. This further raises the ante for international organizations, to promote rural livelihoods in addition to those that may be gained from agriculture.

The idealized AKIS/RD model (Figure 4) purports to integrate agricultural education, research and extension services for capacity building among agricultural producers. In recent documents formulated by the Food and Agriculture Organization of the United Nations, support systems such as marketing and supplies are also included as

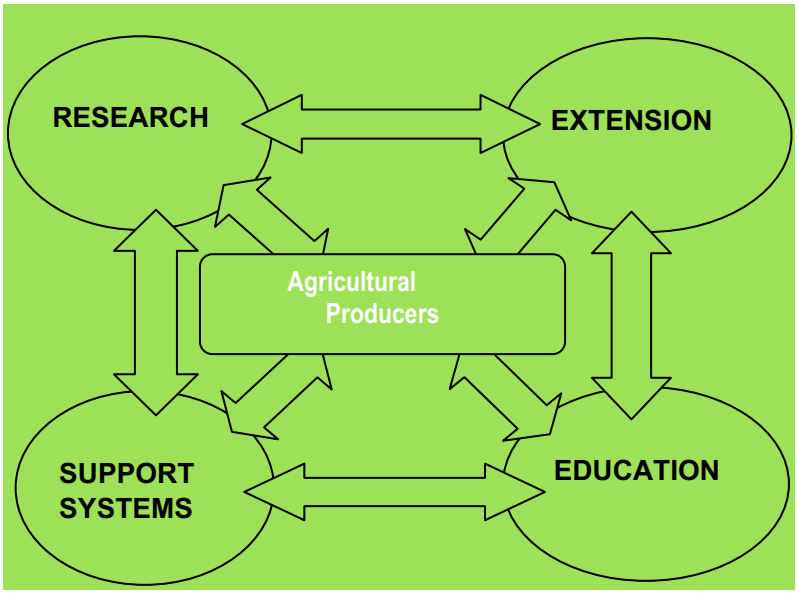
---

[6] For reference, see: http://lnweb18.worldbank.org/ESSD/ardext.nsf/11ByDocName/PublicationsExtensionReformforRuralDevelopment

necessary ingredients in the model. An even more inclusive model of AKIS/RD includes other relevant components, such as: government policy, institutional commitment, and communication systems aside from those used by agricultural extension services.

The AIS model also suggests a complexity of interconnections. In addition to research and extension, the model includes the media, producer organizations, input suppliers, financial institutions, NGOs, etc. All these stakeholders create a spider web and that is the idealized way in which the AIS model is presumed to work.

Figure 4. Idealized AKIS/RD Model



Furthermore, as already mentioned, other frameworks are beginning to be advanced, which address both agricultural and non-agricultural rural development concerns. The RKIS comprehensive view brings forth yet another framework, which calls for either separate or parallel systems of coordinated rural knowledge and information networks aimed at diffusing development information on non-agricultural subjects, i.e. *subjects other than those generally covered by agricultural education, research and extension*, as well as for agricultural development. This RKIS model suggests the advancement of parallel systems of extension and non-extension communication services to serve agricultural as well as non-agricultural clientele.

Impacts will vary from framework to framework depending on the placement of research along the spectrum of R&E activities and the scale of these activities. Depending on its placement along the spectrum of R&E activities, the impact of R&E may be considered an outcome or an input. For example, if the framework stresses research then the outcome or impact would be the successful development of research. If, however, the framework stresses a larger domain, e.g. AKIS/RD or rural development, then research would be an input into that larger purpose. Likewise, depending on the size

of the R&E activity, research and extension may exert minimal or maximum impact on farmers and consequently on policy and/or institutional development.

Review

The challenges cited herein span a variety of issues including the models and norms to be used, the diverse types of programs and possible impacts, the importance of new ideological trends, and the plethora of agricultural development concepts. Some of these challenges will need to be met whatever the final determination about whether or how to assess impacts. Ultimately, evaluations aimed at assessing impacts should at least provide interesting records (and lessons) regarding what programs have achieved, or not, and in certain cases may help to show — not just what programs may have achieved but what they may have prevented from occurring!

COMMENTS BY SURESH BABU ON ASSESSMENT OF CAPACITY-BUILDING ACTIVITIES

Assessing the impact of capacity-building activities is a very challenging task.

In the case of IFPRI, one of the main questions would be how to translate the successful Asian experience of building capacity into Africa. Also, what type of capacity building is best for this region?

In response, I would like to add some of my own views on the impact of capacity strengthening programs. A fundamental question that we should ask is: "capacity strengthening for what?" Capacity strengthening is an activity that involves both building above national institutions and human capacities in the university system, national agricultural research systems, and in public institutions such as the Ministry of Agriculture and Food. How to develop quick capacity for impact and how do we assess such impact on policy changes, technological changes, and institutional changes have been of interest to impact evaluators. One way to assess the capacity of the institutions and the human resources within those institutions is to follow up after the capacity strengthening activities in order to find out who is doing what with the capacities that were built. Are they are doing they job better after the capacity strengthening intervention? Why they are not able to use the skills that were imparted during the capacity building activities.

Capacity strengthening could range from strengthening farmers' organizations and water user associations for managing their activities and programs efficiently, to strengthening the parliament committees on food and agriculture. In between these extremes one would find capacity strengthening of policy analysts which are "fire-fighters" in the ministries, and policy researchers in the academic institutions and policy educators in the university systems.

In assessing capacity one has to distinguish the process impact, which is basically strengthening the capacity of the institutions in order to function better, from the final

impact of how the institutions and the human capacity within them have changed the programs and policies "on the ground".

Capacity assessment could also differ according to the institution that is under consideration. For example, within the university system one might want to know how many courses have been newly introduced based on the capacity strengthening of the university teachers and how the curriculum of these courses and the existing courses have changed because of the new capacity strengthening activities.

Capacity-strengthening assessment also needs to consider the general food policy capacity strengthening versus specific methodology-based capacity strengthening, such as CG modeling. Capacity-strengthening impact would also be identified based on the project-specific impacts versus general capacity building within the national institutions. In developing methods of assessing capacity impacts, it is important to identify the indicators of capacity development, timeframe for capacity assessment and how the capacity should be assessed by getting unbiased evaluation including both internal and external evaluators. All challenging aspects to measure!

DISCUSSION AMONG PARTICIPANTS

James Hanson highlighted that it is very important to clarify how researchers are evaluated at IFPRI: whether their performance is linked to publications/outputs or it also includes the influence of their research on policymaking. This will also be linked to how IFPRI rewards researchers.

Rob van den Berg mentioned that impact is a relative term, and is relative to the ambitions of a project and the initial objectives and expectations of what the project wants to achieve. Comparing the goals/results of research organizations like IFPRI to business organizations, we can see that the latter have much less ambitious objectives. The question is why in development we set much more ambitious goals. Van den Berg also stated that the "logical framework" cannot be applied to policy research activities. This framework is characterized as being static, assumes linear thinking (vs. game theory applications) and causal linkages (e.g., attribution, contribution). The logical framework makes us think that there is a causal link between research and its impacts and that is not always the case.

David Nielsen stated that it is not very clear what IFPRI's IA objectives are. Researchers at IFPRI are not very clear on what IFPRI's role is and which are the areas in which IFPRI can get involved. A big part of what IFPRI does is not easy to observe and evaluate. Therefore, some other, more appropriate indicators should be developed. Also, we should consider how to incorporate intangibles in these indicators.

Ryan responded that it might be a good decision to link impact assessment to the strategies and priorities of IFPRI. We should reflect on how to factor the policy environment into choosing the research countries. IFPRI's new Development Strategy

and Governance Division has strong links to Ethiopia; this might be an example of a bias towards choosing a research country where there is a good policy environments. Additionally, IFPRI should start reflecting on whether to link impact assessment to staff performance assessment. A decision on this question must take into consideration whether different researchers are expected to have different degrees or types of impact, as well as whether their research produces global public goods.

2. NEW AND OLD IFPRI

Session chaired by Rob van den Berg

EX ANTE IMPACT ASSESSMENT: AN EXAMPLE FROM HARVESTPLUS ACTIVITIES IN INDIA — J.V. MEENAKSHI

The biofortification strategy consists of enhancing the micronutrient density of staple crops consumed in the developing world and also to ensure that is bioavailable. Biofortification aims at increasing the proportion of the population that is iron sufficient.

To measure the impact of this program, HarvestPlus will use Disability-Adjusted Life Years (DALYs) as an indicator. DALYs define a unit of health, and they enable: (1) measurement of improvements below the threshold of head count; (2) addition of different disease outcomes, even death; and (3) measurement of benefits with or without ascribing a price.

DALYs are the sum of the years of life lost due to cause-specific mortality and the number of years lived with disability, suitable weighted by severity.

$$DALYs_{lost} = \sum_j T_j M_{ij} \left(\frac{1 - e^{-rL_j}}{r} \right) + \sum_i \sum_j T_j I_{ij} D_{ij} \left(\frac{1 - e^{-rd_{ij}}}{r} \right)$$

Where *j* denotes the target group and *i* the disease, *T* is the size of the target group, *M* the mortality rate, *L* is remaining life expectancy, *r* the discount rate, *I* the incidence rate, $0<D<1$ the disability weight, and *d* the duration of the disease .

Biofortification is expected to reduce the mortality rate (M) and the incidence rate (I).

For iron deficiency, diseases are:

- ✓ Impaired physical activity (moderate & severe iron deficiency anemia, or IDA)
- ✓ Impaired mental development(moderate & severe IDA)
- ✓ Maternal mortality, including stillbirths and child deaths (severe IDA)

The target groups j are:

- ✓ Children aged ≤ 5 years
- ✓ Children aged 6-14 years
- ✓ Women 15+ years
- ✓ Men 15+ years

Some examples of disability weights include:

- ✓ Impaired physical activity from moderate IDA: 0.11 (all target groups)
- ✓ Impaired physical activity from severe IDA: 0.87 (children); 0.90 (adults)

Table 2: Iron Deficiency Anemia

Target Group	Current prevalence rate for	
	Moderate IDA	Severe IDA
Children < 5 years	0.275	0.032
Children age 6-14 years	0.156	0.008
Women 15+ years	0.074	0.010
Men 15+ years	0.037	0.005

In addition to the figures from the previous table, we know that maternal mortality is 540 deaths per 100,000 live births, 5 percent of this figure is assumed to be due to severe IDA. Furthermore, 30 percent of maternal deaths result in stillbirths and 13 percent of surviving and otherwise breastfed infants die.

Calculations show that the number if DALYs lost due to iron deficiency (status quo) is 4 million.

The impact of biofortification can be summarized in the following tables.

Table 3: Impact of Biofortification

	Iron rich rice		Iron rich wheat	
	Pessimistic	Optimistic	Pessimistic	Optimistic
Increase in iron content	50%	300%	5%	10%
Bioavailability	Unchanged			
Share in production	42.5%	85%	47.5%	95%
Full adoption in	10 yrs	7yrs	7 yrs	5yrs

Table 4: Preliminary estimates of DALYs

	DALYs lost due to iron deficiency	DALYs gained through biofortification	Reduction of the burden of iron deficiency
Status quo	4.0 m	none	None
Pessimistic scenario	1.8 m	2.2 m	– 54%
Optimistic scenario	0.4 m	3.5 m	– 89%

Based on author's own calculations

Also important is to consider whether biofortification is cost-effective. Guesstimates of basic R&D, country-specific and maintenance costs suggest that these would translate into less than $1 per DALY saved. This is considered "highly" cost-effective. It appears to be more cost effective than alternative interventions. Even though estimates are preliminary, they exemplify how application of this methodology can help guide resource allocations across micronutrients, crops and varieties, regions, and other interventions.

However, models are only good as the assumptions that underlie them. Therefore, several questions remain: How good are intake data? How valid are assumptions about intra-household food allocations? What do we know about the diets of the very young? What is the nature of the dietary transition? Is there an increasing disconnect between production and consumption decisions? Has there been progress in the coverage of public health programs?

We expect that modeling exercises supplemented with qualitative surveys will provide some answers, although we know that it will also create more questions.

COMMENTS BY GERSHON FEDER

We should take into account what counterfactual (static or dynamic) is being used to compare the program impact. The most suitable for this type of analysis should be dynamic. Related to the counterfactual issue is the question of which alternative approaches are available to achieve the fortification objectives in the absence of the proposed program, and what their costs are. Finally, we should also consider diffusion efforts and how their costs were incorporated into the program.

RETROSPECTIVE NARRATIVES OF OUTCOMES, INFLUENCE, AND IMPACT — MARC J. COHEN

"Statistics = Facts; Stories = The Truth"
— Bangladeshi villager, as reported to IFPRI Research Fellow Kelly Hallman

Institutionalizing impact assessment at IFPRI is key so that researchers see such a process as an integral part of learning and enhancing the quality and cost-effectiveness of their research. With this objective, IFPRI designed two strategies for institutionalization: (1) *ex ante* impact evaluation on new projects, such as that discussed in J. V. Meenakshi's presentation; and (2) group interviews with researchers in which they narrate uses of the research they produce and the influence that research had on the policy decisionmaking process.

In the fall of 2002, IFPRI's Director General asked for a test run of narratives on outcomes, influence, and impact by IFPRI researchers to capture their rich insights and experience and learn from the process. This was the first round of an ongoing effort to assemble narrative material from all research staff on a regular, systematic basis in order to build a database on potential impacts.

Narratives are recollections of research outcomes and of instances where researchers found that their research outputs were influential in the policymaking decision process. One of the objectives of these narratives is, as mentioned, to institutionalize impact assessment at IFPRI and to create a culture of evaluation and self-learning in which researchers are motivated to think about the outcomes and influence of

their research to enhance the Institute's accountability. But ultimately the objective is to contribute to the quality and relevance of the Institute's research, capacity strengthening, and communications.

As used by social scientists, the use of narratives is a research method in which stories present sequences and contexts for a series of events, with a clear beginning, body, and conclusion: What happened? How did it happen? Who did what to make it happen? The use of narratives has some limitations that should be considered when analyzing the information collected with this research method. Narratives can be subjective and open to interpretation; the storyteller may modify or withhold certain information either deliberately or subconsciously—being either too modest or too grandiose; IFPRI researchers may not be fully aware of the range of influence and impact of their own research; and, findings are not easily generalized, despite the value of institutional learning.

Based on the mentioned limitations, narratives work best in conjunction with other approaches. Statistical studies may purchase empirical rigor with context and process, hence narratives are a good complement. A good example of the complementarity of narratives is the analysis of rice policy changes in Viet Nam and the contribution of IFPRI's policy research by Ryan (1999a). In this analysis, Ryan uses an economic surplus model to gauge the value of policy change induced by IFPRI research and also narratives to capture the policy debate around this issue.

For IFPRI's IA activities, narratives will be used mainly for three purposes: (1) as a baseline for institute-wide performance indicators; it should be noted that they will not be used to assess individual performance; (2) to generate important inputs into more in-depth case studies of the impact of IFPRI research, e.g. the in-progress assessment of the gender and intrahousehold research that used some of the narratives from the first round of interviews; and (3) for institutional learning and memory.

Because the policymaking process is complex and research is only one input among others, narratives are not looking to *attribute* impact to a specific piece of research, but they are looking for the *contribution* of research on a certain policy.

Narratives are stored in a database that allows retrieval of inputs for future evaluations of IFPRI research, mainly for impact assessment case studies. These narratives will serve as benchmarks for future evaluations. Moreover, maintaining this information in the database will contribute to creating a corporate memory that is not lost when individual staff members leave. At present, the information in this database can be organized by program, project, division, and narrator. The goal is to combine this database with other existing databases at IFPRI (e.g., donors' database, publications' database), and in this way information will also be retrieved by donor and country, and will include data on outputs.

So far, there have been two rounds of narratives in which most researchers at the postdoctoral level and above have been interviewed in cross-divisional groups. From the

first round of interviews, which took place in late 2002 and early 2003, we can conclude that: (1) researchers familiarized themselves with narrating the use and influence on policy of their research; (2) there was a strong synergy among programs and divisions; and (3) most researchers were reluctant to attribute impact to their research. The second round of interviews began in the fall of 2004 and is still underway. In this round researchers were provided with a guiding questionnaire in advance, to help facilitate good "story telling." Researchers were asked to narrate new stories from 2003, including impacts of past activities that became evident in 2003. In contrast, in the first round researchers were asked to describe any influence regardless of the year, to capture all previous stories.

A preliminary conclusion from the second round of interviews is that claims of influence continue to be modest. In this round of narratives researchers have a better understanding of what this process is about and its objectives, and therefore are more open to share stories. Also, there was a very positive interaction among researchers in group interviews, many times commenting on other researchers' work and influence on policymaking.

The question for discussion in this workshop is how to capture this information on influence in policymaking and outcomes if not through periodic group interviews.

COMMENTS BY JIM RYAN

It is important to emphasize the role of Joachim von Braun, IFPRI's Director General, in advancing this method of collecting information. Von Braun backed this initiative from its first moments and he still believes that is a valid exercise for the Institute.

Narratives encourage researchers to keep an eye on the influence their research has on policymaking. This exercise helps set researchers' minds into thinking beyond just their research.

Validation and corroboration are key in this process. We need to make sure that follow up on the narratives can occur. This process of corroboration will validate the narratives and provide credibility to the stories.

The database where all the narratives are organized and stored is a very useful tool, for several reasons. In particular, when staff move on, and take with them their knowledge of the influence of their research, the database will still contain their narratives.

COMMENTS BY PARTICIPANTS

Feder asserted that, with regard to attribution problems, we have to understand that we are dealing with social science research, and its nature is uncertain.

Hanson pointed that if we assume that the purpose of these narratives is developing a "culture" of evaluation, we should consider the strengths and weaknesses of hiring external reviewers. On the one hand, external reviewers will bring good expertise and objective views, because they are not analyzing their own research. But on the other hand, internal researchers tend to de-link themselves from the assessment thinking that its not their responsibility anymore.

Kym Anderson mentioned that in his evaluation of IFPRI's computable general equilibrium (CGE) modeling he used narratives.

Van den Berg also pointed out that data in a qualitative approach such as narratives are no more uncertain than those in many quantitative approaches. Therefore, the question of verification can be applied to both types of research.

3. BEYOND IFPRI

Session chaired by Susan Offutt

INSIGHTS FROM A CGIAR META-EVALUATION — BRUCE GARDNER

The CGIAR assessment was a review of previous IAs and evaluations. The literature on IA of policy research by the CGIAR system is very scarce, and most of it comes from IFPRI itself. The paper under discussion (Gardner 2003) was made available to participants. It had been prepared as a background contribution to a wider study of the CGIAR (OED 2003).

This evaluation found that IFPRI is good at being close to its original mission of doing policy research in developing countries. But we also noticed that IFPRI has no competitive advantage in projection studies. IFPRI has also gone too far into methodology development, e.g. in CGE modeling. A niche for IFPRI is to study the impact of developed countries' policies on developing countries. This would make use of multi-commodity models, but not necessarily CGE models. IFPRI needs to do more work in the field and less in Washington, DC.

COMMENTS BY MARK ROSEGRANT

The meta-evaluation demonstrates the impact of impact assessment more than the authors may be aware. It should also be noted that many of the recommendations mentioned in the report in fact have been implemented. For example, one of the recommendation states that "The African NARS hoped not only for a sharing of the limelight and of responsibilities but also of resources (especially when African states became paying members of the CGIAR), and this effort proved disappointing." Since then the Africa Challenge program has moved toward remedying this. Although IFPRI is praised for a desire to assist with policy analysis, the report says, "IFPRI needs to review its capacity building and partnership very urgently if it is to help the African NARS to negotiate for themselves in the policy platform, which now includes such diverse themes as globalization, world trade, biotechnology...." IFPRI has responded to these arguments, establishing a large office in Addis Ababa, and also offices in Kampala and Dakar. IFPRI has also institutionalized a number of Africa networks, and had promoted education through helping to develop a collaborative Master of Science Programme in Agricultural and Applied Economics in Eastern and Southern Africa.

Studies show continued high rates of return to agricultural research, but in this report the authors note, "Evidence of recent slowdown in rates of increase in crop yields in many countries may mean that a decline in returns to research is now occurring". In reality, this could easily mean that there has been a decline in overall research funding. Has CGIAR research stayed at the same marginal equilibrium? Why is there no

diminishing marginal return? Might this have quasi-market effects like declining commodity prices that kept policy research at appropriate levels?

The report speculates about private sector taking over much of the research but this is still somewhat unlikely. Is a CGIAR-private sector partnership a feasible transition? While private investment in Asian agriculture is likely to grow, an important role for public investment in the sector will remain. This is because of the incentive problems that discourage private investment in some agricultural research and, also, because of government desire to pursue equity or poverty-alleviation objectives. Agricultural research is in many cases long-term, large-scale, and risky, which means that most firms cannot carry out effective research and institutions may have to be set up on a collective (industry-wide or government) basis to achieve an economically efficient size and scale. The returns on new technologies are often high, but the firm responsible for developing the technology may not be able to appropriate the benefits accruing to the innovation—as in the case of improved open-pollinated rice and wheat varieties. The benefits of agricultural research often accrue to consumers (through reduction in commodity prices resulting from increased supply), rather than to the adopters of the new technology, so social returns may be greater than private returns to research. Appropriate government investments and policy interventions are therefore warranted, especially in areas with relatively low private incentives and relatively high social payoffs.

The report states that "Impact assessment of CGIAR activities… [are] for the World Bank or other donors to use in allocating their global public good budgets." The question would be why are we doing impact assessment? Can impact assessment drive future research priorities and World Bank investments?

The argument on impact assessment is biased, favoring a research with quantifiable returns.

What about introducing markets and competition into setting CG research priorities? What is the role of competitive research/markets for research? Competitive research was, for example, initially proposed by the Steering Committee of the Challenge Program Water and Food, but has not had promising results thus far. It may be the case that aggregate markets are already working, as noted above.

In addition, research already faces market test. At IFPRI, 70 percent of research funding is special projects, which have to be sold on the market.

Should there be more rates of return studies on agricultural research generally? Unless the studies can better differentiate types of research or research on specific varieties and traits, I would argue no. So what then should be areas of future impact assessment?

On capacity building, the World Bank report states: "Attempts to place a value... However, a recent more general review of CGIAR training activities found plenty of evidence that training occurred but no citable results of that training." It is recognized that more capacity-building or -strengthening activities are needed by the NARS. The

International Rice Research Institute's capacity building program generated big networks of rice scientists who generated new germplasm in NARS and became senior policymakers. Surely a well designed impact assessment could capture these impacts.

On natural resource management and the environment, the report seems to equate a lack of existing impact assessments in this area with a finding of no or low impact. "Largely unknown: (1) Rates of return to parts of the expanded CGIAR agenda — fostering biodiversity, countering global warming, social science research — are highly conjectural. Uncertainty of returns would be fine if expanded agenda were costless, but not if it is diverting research resources from activities with higher expected returns." This is far too conservative a test on new areas of research — no one would ever do anything new if that test were followed.

On NARS-CGIAR relationship, can IA help in assessing disaggregated benefits and future roles? But are the future roles related to relative impact or political decisions, and to what extent should CGIAR priorities be driven by NARS and other stakeholder demand? The report notes that India would like the CGIAR to get out of the business of working on environmental issues as global public goods. However, I would argue the exact opposite, that the CG should work more on the environment and natural resources because that is an area that is given low priority by developing country policymakers interested in short-term programs, and by the NARS that depend on those policymakers. More generally, might it not be more productive for CG centers to work counter to demand for at least a subset of their work?

A critical IA of CGIAR systemwide initiatives, Challenge Programs, and other top-down alliances would require a full cost-benefit analysis with careful analysis of transaction costs.

COMMENTS BY PARTICIPANTS

Rivera mentioned that capacity building is more like training without results, and this is because it is not demand-driven.

Van den Berg mentioned that a conservative approach is appropriate when allocating funds. There are areas that need funding regardless of "expected" impact. Furthermore, rate of return of research is a linear way of thinking without thinking of the many complexities of the impact of research.

Hanson mentioned that there is no clear evidence that training works. Van den Berg questioned this statement saying that there is not a clear definition of what is evidence in this case. He argued that assessment did not necessarily need a counterfactual.

ASSESSING DEVELOPMENT DIALOGUES: THE 2020 EXPERIENCE — RAJUL PANDYA-LORCH

The 2020 initiative at IFPRI is mainly an initiator and convener of dialogues, and it also contributes to dialogues. In 1999, Robert Paarlberg conducted an IA. Three main areas were analyzed: whether 2020 had (1) reached the audience; (2) had an impact on the policy thinking of this audience; and (3) catalyzed any new policy action among this audience.

The assessment focused on three main audiences:

1. Researchers and educators, where the assessment found that the 2020 initiative had significant success in reaching them.
2. International policy leaders (initial audience of 2020); 2020 had high success in reaching them, significant success in catalyzing consensus, and noticeable success in catalyzing new policy actions.
3. Policy leaders in the developing world (not an initial target audience); 2020 had significant success in reaching them, noticeable success in catalyzing consensus, and no noticeable success in catalyzing new policy actions.

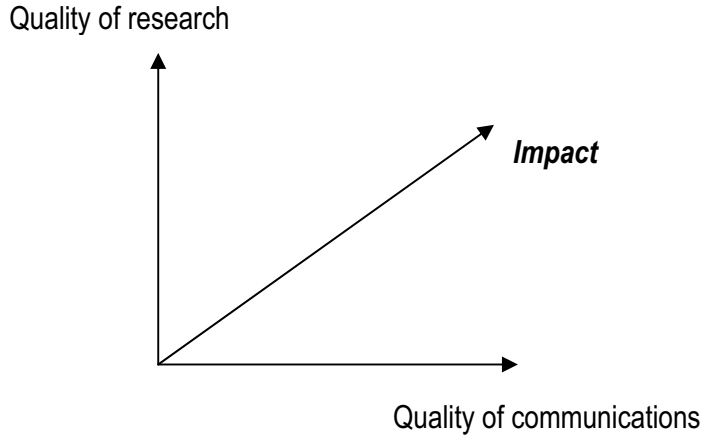
Another assessment of the 2020 activities was a survey of the Bonn Conference (2001) participants conducted by Jim Ryan. The main results of this survey include:

1. There was consensus that the conference was well organized.
2. The content was appreciated, and it provided new and/or surprising information and insights.
3. The expectations were met about the quality of speaker and discussions and also on the wide diversity of topics and stakeholders.
4. There was unhappiness about excessive time devoted to presentations, and participants also expressed preference for breakout sessions.
5. The participants were disappointed about the lack of agreement on clear priorities and action plans.

The 2020 initiative is regularly assessing *what* topics to initiate dialogue on, *which* ongoing or nascent dialogues to contribute to, *how* to contribute to dialogues, w*hen* to engage in/contribute to dialogues, *whom* to include in dialogues, *where* to conduct dialogues, *how* to design dialogues, and *how* to assess dialogues.

COMMENTS BY KLAUS VON GREBMER — A COMMUNICATOR'S POINT OF VIEW

Figure 5: Good Research Needs Good Communications to Have Impact



The results of food policy research are mainly used to: (1) confirm the appropriateness of the actions taken; (2) indicate the actions needed to reduce risks/costs or increase benefits; (3) show ex ante outcomes of alternative policies; (4) show how other policymakers coped with an issue (synthesis); and (5) alert policymakers to major threats.

Policymakers need the right information, in the right form and at the right time. In order to increase their success in reaching policymakers, researchers need to understand the perspective of the policymaker and understand the policy process. Research should (1) be solution oriented; (2) be practical and pragmatic; (3) be persistent; (4) be credible; and (5) quantify the impact of different policy options.

Impact is often two-way: IFPRI needs to look at how multi-stakeholder dialogues impact the Institute.

COMMENTS BY PARTICIPANTS

Ryan pointed that the influence of the spoken word is greater than the influence of the written word. He also questioned the timing for evaluating the 2020 Conferences: right after the conference or after a couple of years.

Rivera mentioned that is always good to ask how participants are going to use information from a conference, and it is important to keep contact with participants.

4. THE EDGE

Session chaired by Marc J. Cohen

THEMATIC IMPACTS: IMPACT ASSESSMENT OF IFPRI'S ECONOMY-WIDE MODELING RESEARCH AND RELATED ACTIVITIES — KYM ANDERSON

There are two main reasons why the economy-wide modeling research and its related activities were assessed in 2003. First, because it was a decade since the Trade and Macroeconomics Division (TMD) at IFPRI began its program of economy-wide modeling. Also, because this division was disbanded in April 2003 and the modelers moved to two other divisions, but nonetheless some questions remained: What is the optimal level of CGE modeling investment at IFPRI? What is the best way to deploy CGE modeling resources within the new structure? What is IFPRI's optimal mix of CGE modeling outputs?

The purpose of this assessment was to provide ex post evaluation of TMD's modeling performance since 1994, bearing the previous questions in mind. That required:

- ✓ Documenting the group's inputs and outputs,
- ✓ Examining the uptake of those outputs,
- ✓ Identifying where that stimulated policy reform, and
- ✓ Assessing the impact of those reforms on well-being in developing countries.

The following is a description of the <u>inputs</u> used in TMD's economy-wide modeling:

- ✓ TMD accounted for 1/10th of IFPRI's budget and total staff, and 1/7th of its researchers in 2002.
- ✓ The budget per researcher averaged just over $100k, comparable with, e.g., the Center of Policy Studies at Monash University in Australia (at just under $90k in 1997–98).
- ✓ 7.4 researchers, 7.8 RAs on average per year (only half were core staff), also similar to CoPS. (Is that above the critical minimum required to be effective/sustainable?)

In general, there was a good mix of complementary skills in the team.

Regarding TMD's economy-wide modeling <u>outputs</u>, there was a respectable rate of publications per researcher, below the IFPRI average for books and journal articles but 30 percent above in book chapters and almost double in discussion papers. If we compare this to the outputs related to IFPRI's International Model for Policy Analysis of Agricultural Commodities and Trade (IMPACT), the total publications per researcher-year of TMD modeling is less than half the rate of the IMPACT project. There is a high quality of publications spread over a wide range of applied policy topics. Also, there are social accounting matrices (SAMs), models, methodologies, and CGE training material

that are state-of-the-art. Numerous training courses had been put on developing countries.

The uptake of TMD's modeling outputs can be summarized as follows:

- ✓ A staggering website download rate in the 14 months to March 2003:
  - (a) 313 per month for its eight most popular discussion papers, compared with 220 for IMPACT's eight most popular publications;
  - (b) An average of 65 per month for all 108 TMD Discussion Papers published since 1994; and
  - (c) Paper No. 75 (published April 2001) had over 22,000 downloads, and another 26 papers had more than 1,000 each.
- ✓ The most popular TMD Discussion Papers had a substantial modelling content and an average of 74 downloads per month, or 58 if No. 75 is excluded from consideration.
- ✓ Those with no modelling content were the least popular, averaging 37 downloads per month.
- ✓ Requests for the new "Standard CGE Model" publication and CD-ROM exceeded 150 per month.
- ✓ TMD received hundreds of requests per year for their modelling expertise:
  - (a) To generate and share SAMs (e.g., with the Global Trade Analysis Project, or GTAP, at Purdue University),
  - (b) To discuss methodological developments,
  - (c) To present model results,
  - (d) To participate in short-term policy missions,
  - (e) To supervise graduate students, and
  - (f) To conduct training courses.

When analysing the impact of TMD's modelling, it can be observed that the attribution problem is acute because TMD covered a full spectrum of basic to applied policy research and its dissemination, plus SAM data compilation and short-term missions and training. The Division also covered all developing country regions, as well as multilateral and regional trade policy issues, and all products (not just food).

To analyze the impact of TMD's modelling approach, a survey of peer researchers and policy stakeholders in developing countries and international institutions was compiled from personal interviews and questionnaires, supplemented with narratives provided by TMD staff.

The following are the main conclusions from the survey responses:

- ✓ Even the least-informed respondents believe economy-wide modelling is 'extremely valuable' for food policy analysis, despite its complexity, because it quantifies effects:
  - (a) of non-food policies on the food sector,
  - (b) of shocks on factor markets and hence income distribution and poverty, and
  - (c) of food sector changes on other sectors.

- ✓ The majority of the more-informed respondents rated both IFPRI and non-IFPRI economy-wide modelling publications as "extremely valuable".
- ✓ Likewise for each of TMD's outputs except multiregional modelling, although a role was seen for IFPRI to combine with other global modellers to generate better projections of the global economy.
- ✓ Strong consensus that TMD's greatest visible contribution to international public goods was in providing SAMs, the Standard CGE Model, and methodologies supporting both.
- ✓ More difficult to attribute policy reform to one research group, but consensus was that economy-wide modelling generally is an effective contributor to the policy process.
- ✓ On the mix of IFPRI's economy-wide modelling products, there was:
  - (a) strong support for more free international public goods (SAMs, models, collaborating with developing countries),
  - (b) very strong support for more work in Africa,
  - (c) also, very strong support for more work on inequality/poverty issues, and
  - (d) less interest in additional short-term policy work and multiregional trade analysis.

Finally, there are three main questions that should be considered by international agencies such as IFPRI:

1. How much should be spent on economy-wide modelling?
   - ✓ To achieve scholarly (hence analytical) reputation versus short-term policy impact.
   - ✓ If not IFPRI, then who will do this on behalf of developing countries (especially least developed countries)? And what would be missed if it is not done (e.g., for Africa in WTO and free-trade arrangement negotiations)?
   - ✓ What is the minimum viable size of a CGE (as distinct from partial equilibrium) modelling group?

2. How should those resources be deployed within the institution?
   - ✓ Single unit gives critical mass, high morale, and analytical reputation.
   - ✓ But risks becoming too academic and interacting too little with other modellers (e.g., IMPACT) and non-modellers, and with the policy community.
   - ✓ Spreading modellers to more-applied divisions may solve the latter.
   - ✓ But… what about modeller critical mass and leadership?

3. What is the institution's optimal mix of economy-wide modelling products?
   - ✓ Methodological versus applied focus.
   - ✓ National versus multiregional (including global) focus.
   - ✓ Stand-alone modelling versus collaboration with other modellers (especially at the global level: GTAP community).
   - ✓ Extent of data compilation (middle- versus low-income countries).
   - ✓ Outreach and training versus research/publishing.

COMMENTS AND IMPLICATIONS FOR THE GOVERNANCE WORK BY REGINA BIRNER

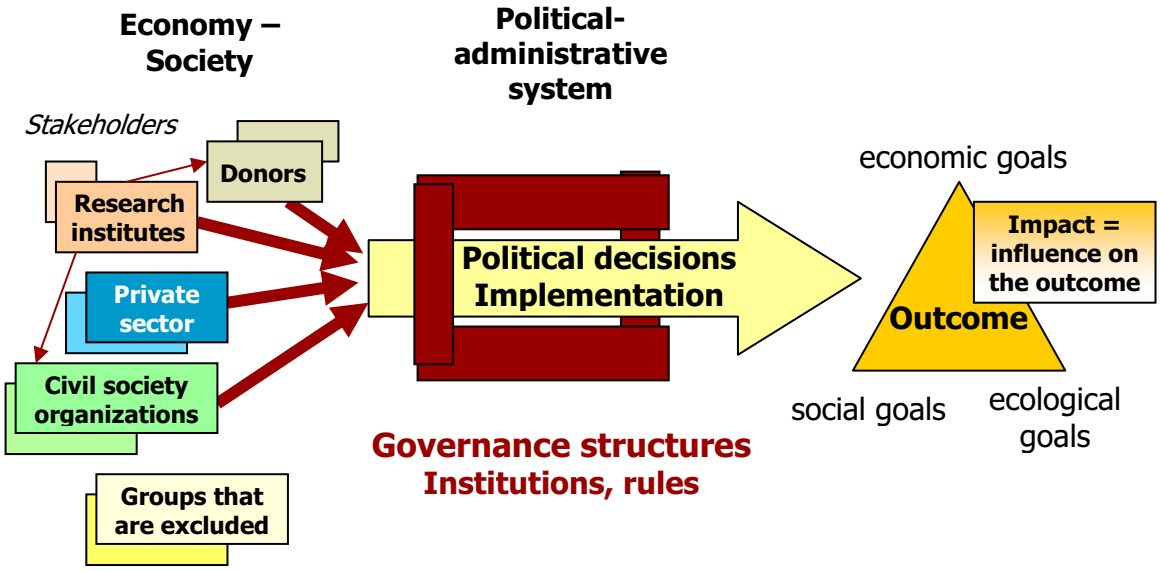
Kym Anderson's presentation is an impressive documentation of the usefulness of IFPRI's economy-wide modeling. It especially allows the analysis of a broad range of policy issues and it even takes distributional aspects into account.

Regarding the uptake of IFPRI's economy-wide modeling outputs, the most important factors are the high number of website downloads and the high-level officials/institutions using it.

The challenge of assessing the impact of economy-wide modeling can be found in the methodologies used. The use of a survey gives an interesting insight on perceived advantages/disadvantages, but the question remains on how representative the results are. Narratives give interesting insights regarding persons/institutions using the outputs. The question is what the underlying model of the policy process is and whether the potential of this method is fully used.

The implications for governance research can be summarized with the following figure.

Figure 6: Model to structure governance research



An analysis of the policy process shows:

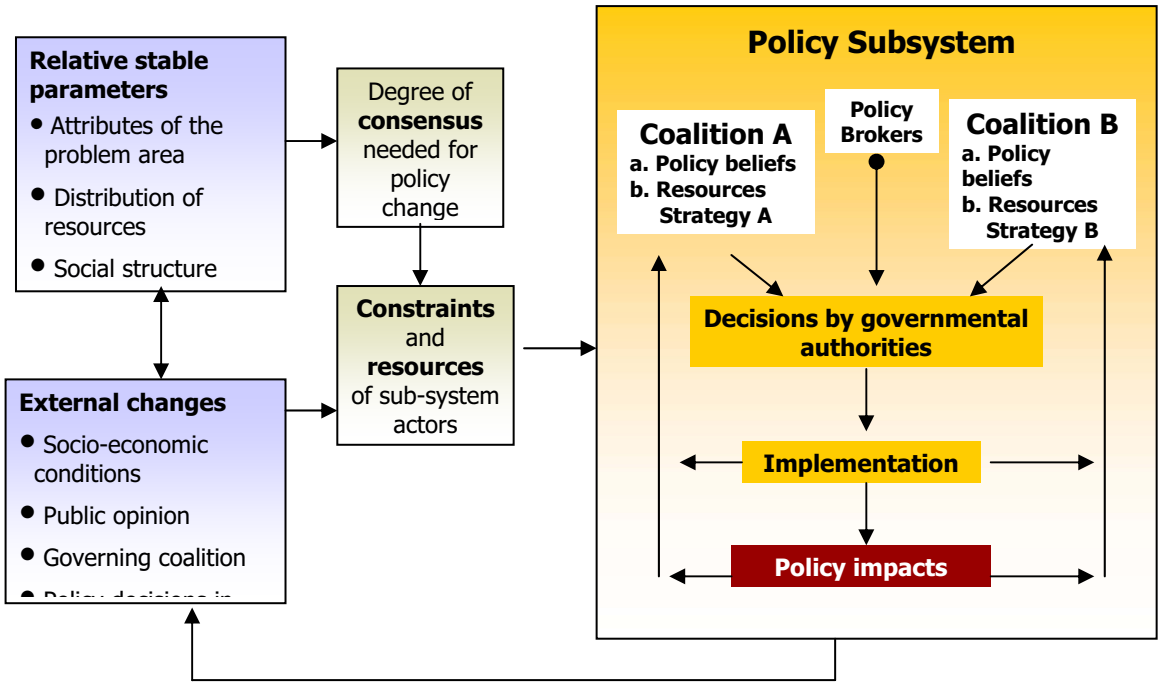
1. Advocacy coalitions (proponents, opponents)
   - ✓ Different actors (agency members, lobbyists, scientists, etc.)
   - ✓ Shared interests and beliefs (core beliefs, policy beliefs)
   - ✓ Resources and strategies to create political capital
   - ✓ Understanding and trust in model results?

2. Timing: Windows of opportunity
   - ✓ Policy options may be well known for long periods
   - ✓ Policy changes if problem situation and/or political conditions create a window of opportunity
   - ✓ Political entrepreneurs and policy brokers important
3. Policy-oriented learning across coalitions
   - ✓ Role of accepted quantitative data
   - ✓ Role of multi-stakeholder platforms and participatory policymaking (example: Poverty Reduction Strategy Processes)

Among the most interesting results, we can see that most respondents find "extremely valuable" the effects of non-food policies. The most frequently quoted disadvantages include that it was "too difficult to communicate results to stakeholders" and that is "too complicated for our staff to use". This might be related to issues of training and policy communications. Finally, when thinking about where IFPRI should do more within its current budget, it is recommended that the money be used to:

- ✓ "Write up CGE results in non-technical policy papers/briefs"
- ✓ "Offer simpler models to assist intuition"
- ✓ "Focus on national trade issues in Africa"
- ✓ "Collaborate with researchers in developing countries"

Figure 7: Analytical model of the policy process — Sabatier's Advocacy Coalition Framework (1988)



COMMENTS BY PARTICIPANTS

Xinshen Diao mentioned that CGE models are without a doubt useful tools that are used in many other organizations. Currently, several divisions at IFPRI use CGE modeling as one methodology among many. Currently, CGE models are linked to household data and spatial data, which boost their power.

Gardner mentioned that it would be very valuable to elaborate on the CGE approach as compared to other tools. CGE is a problem in many cases because it does not carry answers to what it wants to answer. This modeling tool is not the best tool for many areas, and it is also weak in looking at the short and long run.

Hazell pointed that IFPRI has not given up on CGE models, but has scaled back research on the methodology itself. IFPRI has intensified the use and applications of this tool.

CHALLENGES OF RANDOMIZED AND QUASI-EXPERIMENTAL APPROACHES — MICHAEL KREMER

Two examples of randomized evaluations:

- ✓ Randomized evaluations in education
- ✓ Randomized evaluations of agricultural extension programs

Lessons from randomized approaches:

1. *Randomized evaluations are often feasible:* Randomized evaluations are labor intensive and costly, but no more so than other data collection activities. Political economy concerns over randomized evaluations may sometimes make it difficult not to implement a program in the entire population. However, these concerns can be tackled at several levels: financial constraints often necessitate phasing-in programs over time, and randomization may actually be the fairest way of determining the order of phase-in.

2. *Retrospective evaluations often fail to match results of randomized evaluations:* Estimates from prospective randomized evaluations can often be quite different from the effects estimated in a retrospective framework, suggesting that omitted variable bias is a serious concern.

3. *NGOs are well-suited to conduct randomized evaluations in collaboration with academics and external funders:* Governments are far from the only possible outlets through which to organize randomized evaluations. Unlike governments, nongovernmental organizations (NGOs) are not expected to serve entire populations. Financial and administrative constraints often lead NGOs to phase in programs over time, and randomization will often be the fairest way of determining the order of phase-in. In general, given that accurate estimates of program effects are international public goods, randomized evaluations should be financed internationally.

4. *Costs can be reduced and comparability enhanced by conducting a series of evaluations in the same area:* Once staffs are trained, they can work on multiple projects. Since data collection is the most costly element of these evaluations, crosscutting the sample can also dramatically reduce costs.

5. *Randomized evaluations have a number of limitations, but many of these also apply to other techniques:* Randomized methods often allow for easier correction for these limitations than do non-randomized methods.

   Sample selection problems could arise if factors other than random assignment influence program allocation. For example, parents may move their children from a class (or a school) without the program to a school with the program. Conversely, individuals allocated to a treatment group may not receive the treatment (for example, because they decide not to take up the program). Even if randomized methods have been employed and the intended allocation of the program was random, the actual allocation may not be. This problem can be addressed through "intention to treat (ITT)" methods or by using random assignment as an instrumental variable for actual assignment.

   A second issue affecting both randomized and non-randomized evaluations is differential attrition in the treatment and the comparison groups: those who participate in the program may be less likely to move or otherwise drop out of the sample that those who do not.

   Finally, programs may create spillover effects on people who have themselves not been treated. If such spillovers are global, identification of total program impacts will be problematic with any methodology. However, if such spillovers are local, then randomization at the level of groups can allow estimation of the total program effect within groups and can generate sufficient variation in local treatment density to measure spillovers across groups.

   In summary, while randomized evaluation is not a bulletproof strategy, the potential biases are well-known and can often be corrected. This stands in contrast to biases of most other types of studies, where the bias due to the non-random treatment assignments often cannot be signed nor estimated.

6. *Publication bias likely huge:* There is a natural tendency for positive results to receive a large amount of publicity: agencies that implement programs seek publicity for their successful projects, and academics are much more interested in and able to publish positive results than modest or insignificant results. However, clearly many programs fail, and publication bias will be substantial if positive results are much more likely to be published.

   It is important to put institutions in place to ensure that negative results are disseminated. Such a system is already in place for medical trial results, and creating a similar system for documenting evaluations of social programs would help to alleviate the problem of publication bias.

7. *Randomized evaluations can shed light on larger and more general issues:* Without a theory of why a program has the effect that it has, generalizing from

one well-executed randomized evaluation may be unwarranted; however, similar issues of generalizability arise no matter what evaluation technique is being used. One way to learn about generalizability is to encourage adapted replications of randomized evaluations in key domains of interest in several different settings. While it will always be possible that one program that was unsuccessful in one context would have been successful in another, adapted replications, guided by a theory of why the program was effective, will go a long way toward alleviating this concern. This is one area where international organizations, which are already present in most countries, can play a key role.

For IFPRI this methodology presents a valuable opportunity. Randomized evaluations can play a valuable role in IFPRI's mission.

- ✓ PROGRESA
- ✓ Not applicable everywhere, but more can be done
- ✓ Agriculture, extension
- ✓ Be pro-active

COMMENTS BY EMMANUEL SKOUFIAS

IFPRI's evaluation of PROGRESA (Mexico's Program for Education, Health, and Food) had a big impact. Previous IFPRI research on intrahousehold allocation influenced the way the government of Mexico designed the program. It continued even after the administration changed. Furthermore, IFPRI was able to build capacity in Mexico.

The fact that most IFPRI evaluation is "hired" creates risks and decreases the objectivity of the evaluation. The Mexican government paid IFPRI directly to evaluate PROGRESA, and this will put a big constraint on the evaluators. If IFPRI wants to do these evaluations as public goods, then it should reconsider the sources of funding.

5. WRAP UP

Session chaired by Jim Ryan

Jock Anderson reaffirmed the objectives of the workshop and also brought back to the table the set of questions that were introduced at the opening of the workshop.

Gardner expressed caution on the use of narratives; he suggested increased use of randomized approaches instead. Hanson responded that narratives provide understanding of the impact on people and are a very good complement to other approaches. They are a valid source of information. Birner also stated that narratives are a very useful tool. It is a cost-effective method where researchers have data on their minds. She also mentioned that it is not a good idea to link them to performance appraisal (moral hazard). Garrett added that narratives are good for institutional learning. At some point, researchers have to start thinking about the impact of their research.

Referring to the use of randomized experiments, Rivera acknowledged that there is a broad range of impact assessment methods that can fit IFPRI's evaluations. Garrett reaffirmed this proposition, stating that IFPRI has to look at the various evaluation methods in different fields, and it would be desirable to move towards using a pool of additional methodologies.

Anderson noted that IFPRI is in the process of combining the information from the first and second round of narratives. This information will be included in an already existing database, and after the database is assessed for its functionality it will be decided whether effort using this method will be sustained. IFPRI will continue to carry out case studies of its impact.

On *ex ante* impact evaluation, Gardner suggested that it would be useful to look at how the United States government approaches evaluation. Each and every program has *ex ante* evaluation and *ex post* evaluation. It should also be noted that *ex ante* evaluation is in many cases more important than *ex post*. Ryan said that the CGIAR Science Council stresses the importance of looking at projected outcomes and that there will be more emphasis on ex ante evaluations.

Anderson closed the session by offering his tentative answers to the questions he had posed.

Q 1: (a) How well have the IA objectives been achieved?

The objectives are still relevant but have thus far only been partially achieved. Additional objectives should not be added until achievement of the present set is more complete.

Q 2: Are further case studies justified?

Yes, for all the good reasons discussed as well as strong demand for them by the Science Council.

Q 3: Is the narratives instrument a worthy approach or should it be modified?

The controversial issue was not resolved in workshop discussion, with a wide range of opinion being expressed. The issue should be reexamined once the second round is completed.

Q 4: Should there be a wider IFPRI effort on ex ante studies?

In principle, Yes. Presently, most program formulation is done without explicit *ex ante* analysis but there is clearly much informal consideration of ultimate impact of new research activities. The mentioned pressures from the Science Council will inevitably lead to more focus on *ex ante* methods in portfolio choice although the costs of such work are non-trivial.

Q 5: Are further such general workshops justified?

Yes, ideally to review critically the application of novel methods to important themes in social science research, which of course, in principle, covers all of the IFPRI and much of the CGIAR portfolio.

REFERENCES

Alston, J. M., C. Chang-Kang, M. C. Marra, P. G. Pardey, and T. J. Wyatt. 2000. A meta-analysis of rates of return to agricultural R&D. Ex pede herculem? Research Report No. 113. Washington, DC: International Food Policy Research Institute.

Alwang, J., and V. Puhazhendhi. 2002. The impact of the International Food Policy Research Institute's research program on rural finance policies for food security for the poor. Impact Assessment Discussion Paper No. 16. Washington, DC: International Food Policy Research Institute.

Anderson, K. 2003. Impact assessment of IFPRI's research and related activities based on economywide modeling. Impact Assessment Discussion Paper No. 21. Washington, DC: International Food Policy Research Institute.

Babu, S. C., and G. Mthindi. 1995. Costs and benefits of informed food policy decisions: A case study of food security and nutrition monitoring in Malawi. *Quarterly Journal of International Agriculture* 34 (3): 292–308.

Balzer, G., and U. Nagel. 2001. Logframe based impact monitoring within the CGIAR system. Annex 4 Background Papers. In *TAC Secretariat, The Future of Impact Assessment in the CGIAR: Needs, Constraints, and Options. Proceedings of a Workshop Organized by the Standing Panel on Impact Assessment (SPIA) of the Technical Advisory Committee (TAC), 3–5 May 2000*, pp. 54–78. Rome: Food and Agriculture Organization of the United Nations.

Campbell, H., and D. Squires. 1998. The role of research in fisheries management: Conserving dolphins in the Eastern Tropical Pacific and exploiting Southern Bluefin Tuna in the Southern Ocean. In *Closing the loop: From research on natural resources to policy change*, eds. S. R. Tabor and D. C. Faber, Chapter 6 pp. 52–87. Policy Management Report No. 8. The Hague, Netherlands: European Centre for Development Policy Management and the International Service for National Agricultural Research.

Dinello, N., and L. Squire. 2002. Survey of policymakers. Bridging Research and Policy Global Research Project, Global Development Network (Draft).

Duflo, E., and M. Kremer. 2003. Use of randomization in the evaluation of development effectiveness. Paper prepared for the World Bank Operations Evaluation Department (OED) Conference on Evaluation and Development Effectiveness in Washington, DC, 15–16 July.

Easterling, D. 2000. Using outcome evaluation to guide grant making: Theory, reality, and possibilities. *Nonprofit and Voluntary Sector Quarterly* 29 (3): 482–486.

Gardner, B. 2003. Global public goods from the CGIAR: Impact assessment. Thematic working paper for the CGIAR at 31: An independent meta-evaluation of the Consultative Group on International Agricultural Research. OED, World Bank, Washington, DC.

Gillespie, S., M. McLachlan, and R. Shrimpton, eds. 2001. Turning the tide: The new challenge for nutrition. A synthesis of the findings from the World Bank-UNICEF Nutrition Assessment (Draft).

Haddad, L., and D. Pelletier. 2003. The nutrition policy process: A proposal for a global and regional research program (GRP). Consultation document for discussion and comment. International Food Policy Research Institute, Washington, DC (Draft).

Horton, D., and R. Mackay. 2003. Using evaluation to enhance institutional learning and change: Recent experiences with agricultural research and development. Paper presented at SPIA/CIMMYT conference on "Why has impact assessment research made more of a difference?," San Jose, Costa Rica, 4–7 February 2002.

Jackson, C. 2005. Strengthening food policy through gender and intrahousehold analysis: Impact assessment of IFPRI multicountry research. Impact Assessment Discussion Paper No. 23. Washington, DC: International Food Policy Research Institute.

Mackay, R., and D. Horton. 2003. Institutional learning and change in the CGIAR: Summary record of the workshop held at IFPRI, Washington, DC, February 4–6, 2003. Impact Assessment Discussion Paper No. 18. Washington, DC: International Food Policy Research Institute.

OED. 2003. The CGIAR at 31: An independent meta-evaluation of the Consultative Group on International Agricultural Research. Operations Evaluation Department, World Bank, Washington, DC.

Pardey, P. G., and J. E. Christian. 2002. The production and diffusion of policy knowledge: A bibliometric evaluation of the International Food Policy Research Institute. Impact Assessment Discussion Paper No. 14. Washington, DC: International Food Policy Research Institute.

Rodrik, D. 1996. Understanding economic policy reform. *Journal of Economic Literature* 34 (1): 9–41.

Ryan, J. G. 1999a. Assessing the impact of rice policy changes in Vietnam and the contribution of policy research. Impact Assessment Discussion Paper No. 8. Washington, DC: International Food Policy Research Institute.

_____. 1999b. Assessing the impact of policy research and capacity building by IFPRI in Malawi. Impact Assessment Discussion Paper No. 11. Washington, DC: International Food Policy Research Institute.

_____. 2002. Synthesis report of workshop on assessing the impact of policy-oriented social science research in Scheveningen, The Netherlands, November 12–13, 2001. Impact Assessment Discussion Paper No. 15. Washington, DC: International Food Policy Research Institute.

_____. 2003. Evaluating the impact of agricultural projection modeling using the 'IMPACT' framework. Impact Assessment Discussion Paper No. 17. Washington, DC: International Food Policy Research Institute.

Ryan, J. G., and J. Garrett. 2003. The impact of economic policy research: Lessons on attribution and evaluation from IFPRI. Impact Assessment Discussion paper No. 20. Washington, DC: International Food Policy Research Institute.

Sabatier, Paul. 1988. An advocacy coalition framework of policy change and the role of policy-oriented learning therein. *Policy Sciences* 21: 128–168.

Scott, A. 2000. The dissemination of the results of environmental research: A scoping report for the European Environmental Agency. Expert's Corner Series, Science Policy Research Unit, University of Sussex, October.

Smith, D. R., with A. Sutherland. 2002. Institutionalizing impact orientation: Building a performance management approach that enhances the impact orientation of research organizations. Chatham, UK: Natural Resources Institute.

Stiglitz, J. 1999. Scan globally, reinvent locally: Knowledge infrastructure and the localization of knowledge. Keynote address for the First Global Development Network Conference, held in Bonn, Germany, December.

Springer-Heinze, A., F. Hartwich, J. S. Henderson, D. Horton, and I. Minde. 2002. Impact pathway analysis: An approach to strengthening the impact orientation of agricultural research. Paper presented at SPIA/CIMMYT conference on "Why has impact assessment research made more of a difference?," San Jose, Costa Rica, 4–7 February.

Sutton, R. 1999. The policy process: An overview. Working Paper No. 118 (August). UK: Overseas Development Institute.

TAC (Technical Advisory Committee) Secretariat. 2001. Summary of the sessions. In *The future of impact assessment in the CGIAR: Needs, constraints and options.* Proceedings of a workshop organized by the Standing Panel on Impact Assessment (SPIA) of the Technical Advisory Committee (TAC), 3-5 May 2000, pp. 7–27. Rome: FAO.

Annex 1: Workshop Program

08:00 – 08:30	Breakfast	
08:30 – 10:15	**Session 1 — State of Play** (Chair: Walt Armbruster)	
	Ambitions for the Workshop	*Jock Anderson*
	Lessons on Impact Attribution and Evaluation	*Jim Ryan*
	Comment	*James Garrett*
	Challenges of Assessing Capacity Building	*Bill Rivera*
	Comment	*Suresh Babu*
10:15 – 10:30	Break	
10:30 – 12:00	**Session 2 — New and Old in IFPRI** (Chair: Rob van den Berg)	
	Ex Ante Impact Assessment: An Example from HarvestPlus Activities in India	*J. V. Meenakshi*
	Comment	*Gershon Feder*
	Narratives in Impact Assessment?	*Marc J. Cohen*
	Comment	*Jim Ryan*
12:00 – 13:00	Lunch	
13:00 – 15:15	**Session 3 — Beyond IFPRI** (Chair: Susan Offutt)	
	Insights from a CGIAR Meta-Evaluation	*Bruce Gardner*
	Comment	*Mark Rosegrant*
	Challenges of Assessing Development Dialogue	*Rajul Pandya-Lorch*
	Comment	*Klaus von Grebmer*
15:15 – 15:30	Break	
15:30 – 17:15	**Session 4 — The Edge** (Chair: Marc J. Cohen)	
	Thematic Impacts: Experience on CGE Modeling	*Kym Anderson*
	Comments and Implications for Governance Work	*Regina Birner*
	Challenges of Randomized and Quasi-Experimental Approaches	*Michael Kremer*
	Comment	*Emmanuel Skoufias*
17:15 – 17:30	**Wrap-up** (Chair: Jim Ryan)	
	Answers to the Qs Posed?	*Jock Anderson*
	Discussion	

Annex 2: Workshop Participants

Name	Affiliation
Gary Alex	USAID
Jock Anderson	IFPRI
Kym Anderson	World Bank
Walt Armbruster	Farm Foundation
Suresh Babu	IFPRI
Todd Benson	IFPRI
Regina Birner	IFPRI
Marc J. Cohen	IFPRI
Dana Dalrymple	USAID
Xinshen Diao	IFPRI
Gershon Feder	World Bank
Bruce Gardner	University of Maryland
James Garrett	IFPRI
Paul Gibson	Economic Research Service, USDA
Dan Gilligan	IFPRI
James Hanson	University of Maryland
Peter Hazell	IFPRI
Maria Iskandarani	Consultative Group on International Agricultural Research Secretariat
Michael Kremer	Harvard University/ Brookings Institution
J.V. Meenakshi	IFPRI, HarvestPlus
David Nielson	World Bank
Susan Offutt	Economic Research Service, USDA
Rajul Pandya-Lorch	IFPRI, 2020 Initiative
Bill Rivera	University of Maryland
Marc Rockmore	IFPRI
Mark Rosegrant	IFPRI
Jim Ryan	Australian National University
David Schimmelpfenning	Economic Research Service, USDA
Manohar Sharma	IFPRI
Ken Simler	IFPRI
Emmanuel Skoufias	World Bank
Rob van den Berg	Global Environmental Facility
Klaus von Grebmer	IFPRI

Jock R. Anderson is a Visiting Senior Research Fellow and coordinates Impact Assessment at the International Food Policy Research Institute (IFPRI).

María Soledad Bos is a Senior Research Assistant in the Director General's Office at IFPRI.

Marc J. Cohen is a Research Fellow in the Food Consumption and Nutrition Division of IFPRI.

IMPACT ASSESSMENT DISCUSSION PAPERS

---

25. **Impact Assessment of Food Policy Research: A Stocktaking Workshop — Synthesis Report,** *by Jock Anderson, Maria Soledad Bos, and Marc J. Cohen (December 2005)*
24. **Regional Policy Networks: IFPRI's Experience with Decentralization**, *by Robert Paarlberg (April 2005)*
23. **Strengthening Food Policy Through Gender and Intrahousehold Analysis: Impact Assessment of IFPRI Multicountry Research**, *by Cecile Jackson (April 2005)*
22. **The Impact of the Food-for-Education Program in Bangladesh on Schooling Outcomes and Earnings and the Contribution of IFPRI Research**, *by James G. Ryan and Xin Meng (November 2004)*
21. **Impact Assessment of IFPRI's Research and Related Activities Based on Economywide Modeling,** *by Kym Anderson (December 2003)*
20. **The Impact of Economic Policy Research: Lessons on Attribution and Evaluation from IFPRI,** *by James G. Ryan and James L. Garrett (October 2003)*
19. **Impacts of IFPRI/ICARDA Policy and Property Rights Research on the Mashreq and Maghreb Project,** *by John H. Sanders and Hassan Serghini (October 2003)*
18. **Institutional Learning and Change in the CGIAR: Summary Record of the Workshop Held at IFPRI, Washington, DC, February 4-6, 2003,** *by Ronald Mackay and Douglas Horton (October 2003)*
17. **Evaluating the Impact of Agricultural Projection Modeling Using the IMPACT Framework**, *by James G. Ryan (February 2003)*
16. **The Impact of the International Food Policy Research Institute's Research Program on Rural Finance Policies for Food Security for the Poor**, *by Jeffrey Alwang and V. Puhazhendhi (December 2002)*
15. **Synthesis Report of Workshop on Assessing the Impact of Policy-oriented Social Science Research in Scheveningen, The Netherlands November 12–13, 2001**, *by James G. Ryan (March 2002)*
14. **The Production and Diffusion of Policy Knowledge: A Bibliometric Evaluation of the International Food Policy Research Institute**, *by Philip G. Pardey and Jason E. Christian (January 2002)*

IMPACT ASSESSMENT DISCUSSION PAPERS

---

13. **Impact of IFPRI's Policy Research on Resource Allocation and Food Security in Bangladesh**, *by Suresh Babu (February 2000)*
12. **A Review of Food Subsidy Research at IFPRI**, *by Curtis Farrar (January 2000)*
11. **Assessing the Impact of Policy Research and Capacity Building by IFPRI in Malawi**, *by James G. Ryan (December 1999)*
10. **External Impact Assessment of IFPRI's 2020 Vision for Food, Agriculture, and the Environment Initiative**, *by Robert Paarlberg (June 1999)*
9. **Returns to Policy-Related Social Science Research in Agriculture**, *by Bruce L. Gardner (May 1999)*
8. **Assessing the Impact of Rice Policy Changes in Viet Nam and the Contribution of Policy Research**, *by James G. Ryan (January 1999)*
7. **The Value of Economic Research**, *by David Zilberman and Amir Heiman (January 1999)*
6. **Policy for Plenty: Measuring the Benefits of Policy-Oriented Social Science Research**, *by George W. Norton and Jeffrey Alwang (December 1998)*
5. **Some Useful Methods for Measuring the Benefits of Social Science Research**, *by Henry E. Kilpatrick, Jr. (October 1998)*
4. **Adding Value through Policy-Oriented Research: Reflections of a Scholar-Practitioner**, *by C. Peter Timmer (October 1998)*
3. **A Proposal for Measuring the Benefits of Policy-Oriented Social Science Research**, *by Donghyun Park (August 1998)*
2. **Measuring the Benefits of Social Science Research**, *by Vincent H. Smith (July 1998)*
1. **IFPRI and the Abolition of the Wheat Flour Ration Shops in Pakistan: A Case-Study on Policymaking and the Use and Impact of Research**, *by Yassir Islam and James L. Garrett (December 1997)*