SCIENCE COUNCIL
CGIAR

Strategic Guidance for *Ex Post* Impact Assessment of Agricultural Research

July 2008

CONSULTATIVE GROUP ON INTERNATIONAL AGRICULTURAL RESEARCH

Editing, layout and proofreading: Green Ink Publishing Services Ltd, UK (www.greenink.co.uk)

Consultative Group on International Agricultural Research
SCIENCE COUNCIL

Strategic Guidance for *Ex Post* Impact Assessment of Agricultural Research

T. Walker, M. Maredia, T. Kelley, R. La Rovere,
D. Templeton, G. Thiele, and B. Douthwaite

Prepared for the Standing Panel on Impact Assessment
CGIAR Science Council

July 2008

Correct citation:
Walker T., Maredia M., Kelley T., La Rovere R., Templeton D., Thiele G., and Douthwaite B. 2008. Strategic Guidance for *Ex Post* Impact Assessment of Agricultural Research. Report prepared for the Standing Panel on Impact Assessment, CGIAR Science Council. Science Council Secretariat: Rome, Italy.

Contents

Tables and Figures

Boxes

Acronyms

ACIAR	Australian Centre for International Agricultural Research
BNARDER	Benue Agricultural and Rural Development Project
CCER	center-commissioned external review
CG-centers	centers of the Consultative Group on International Agricultural Research
CGIAR	Consultative Group on International Agricultural Research
CIAT	Centro Internacional de Agriculture Tropical (International Center for Tropical Agriculture)
CIFOR	Centre for International Forestry Research
CIMMYT	Centro Internacional de Mejoramiento de Maíz y Trigo (International Maize and Wheat Improvement Center)
CIP	Centro Internacional de la Papa (International Potato Center)
DALYs	disability-adjusted years
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária (Brazilian Agricultural Research Corporation)
epIA	*ex post* impact assessment
EPMR	external program management reviews
FAO	Food and Agriculture Organization of the United Nations
FFE	Food for Education
ICARDA	International Center for Agricultural Research in the Dry Areas
IARC	international agricultural research center
ICRAF	World Agroforestry Centre (formerly International Centre for Research in Agroforestry)
ICRISAT	International Crops Research Institute for the Semi-Arid Tropics
IDRC	International Development Research Centre
IFPRI	International Food Policy Research Institute
IITA	International Institute of Tropical Agriculture
ILRI	International Livestock Research Institute
IPM	integrated pest management
IRRI	International Rice Research Institute
IWMI	International Water Management Institute
LSMS	living standards measurement study (survey)
MDGs	Millennium Development Goals
MTP	Medium Term Plans (the CGIAR's)
NARES	national agricultural research and extension systems
NARS	national agricultural research systems
NGO	non-governmental organization
NRM	natural resource management
NPV	net present value
PIPA	participatory impact pathways analysis
POR	policy-oriented research
R&D	research and development
RAPID	Research and Policy in Development
rBST	bovine somatotrophin
SPIA	Standing Panel on Impact Assessment
WHO	World Health Organization

Acknowledgements

Hans Gregersen played an important role in identifying the need for these guidelines and in shaping the final product. James Stevenson assisted in developing the outline on which these guidelines are based. Participants at a Standing Panel on Impact Assessment (SPIA) workshop in Nairobi in 2006 discussed the outline in particular and the guidelines in general. Jim Ryan made extensive comments on the first and final drafts of the manuscript and contributed material to the assessment of policy-oriented research – one of the subjects of these guidelines. Jock Anderson, Kwesi Atta-Krah, Flavio Avila, John Dixon, Keith Fuglie, Meredith Girodano, Diakalia Sanogo, Meredith Soule, and Jonathan Wadsworth commented on a later draft of these guidelines. Jeff Davis, Bruce Gardner, Hans Gregersen and David Raitzer reviewed a final draft of the manuscript. Princess Ferguson edited all the drafts and Green Ink did the final editing of this report for publication. We are indebted to all of the above for the time they took to make these guidelines a reality.

Foreword

The Science Council's Standing Panel on Impact Assessment (SPIA) has for some time been aware of the need for a document which provides strategic guidance to impact assessment practitioners and research managers for planning, conducting, and managing *ex post* impact assessments (epIA). We are very pleased to have been able to engage a number of experienced practitioners to distill from their own and other's experiences an authoritative statement which SPIA substantially endorses. Although this document does represent the authors' views, SPIA was involved in crafting the terms of reference on which it was based, interacted with the authors throughout, and has reviewed all the major drafts. Indeed one of the authors is a member of SPIA and another is secretary to it. Hence SPIA should be regarded as the handmaiden and midwife to these guidelines.

These guidelines do not attempt to indicate what best practices are but rather looks at the options and discusses their pros and cons as a prelude to offering 'good practice' advice at the end of each section. This is in recognition of the fact that it is not possible to be definitive in many instances. A number of 'how to' manuals are available that complement this document by providing specific field guidelines to practitioners.

The content of the document is heavily influenced by the long history and extensive experience of the Consultative Group on International Agricultural Research (CGIAR) and national agricultural research systems (NARS) in assessing the impact of research for crop genetic improvement research. Recent efforts to assess the impacts of policy-oriented research (POR) have revealed some key strategic issues around this theme which these guidelines address, although in a limited fashion. This is also the case with the epIA of natural resource management (NRM) research. There is a need for much greater attention to epIA methodologies and case studies of policy-oriented research, of NRM research, and also of training and capacity building to give a more balanced coverage across these issues. Future editions of these guidelines should include more specific references to these neglected thematic areas.

The primary emphasis in this document is on epIA from the accountability perspective. It also refers to learning from impact assessments, although the limitations of epIA in this regard are such that SPIA contends there are other monitoring and evaluation instruments that are more effective for learning than epIA, as it is defined here. There may well be opportunities to enhance the value of epIA for learning and other monitoring and evaluation activities in the future and SPIA is initiating a study this year to explore this.

The document suggests two phases for carrying out epIA. Stage I focuses (although not exclusively) on economic rate of return approaches, whilst Stage II examines multi-dimensional impacts further along the impact pathway to try and capture poverty, social, and environmental impacts. The former is more straightforward than the latter. Stage II epIA requires skills that many research institutions do not possess. Needless to say, Stage II epIAs are becoming mandatory for investors in international agricultural research. Stage I epIAs complement Stage II epIAs, although they are increasingly being recognized as being insufficient to satisfy donor demands for accountability. These realities raise many methodological and strategic issues, and these are discussed in this document. Important among them is the dearth of environmental indicators for epIAs to use. These are not only needed for NRM research but also for multi-dimensional epIAs covering other important themes including crop genetic improvement and policy-oriented research. The link between research outputs and decision-making outcomes has so far been weak in policy-oriented research epIAs. Alternative poverty indicators receive significant attention in this document; but to date few epIA studies have adequately documented impact on alleviating poverty of research. Much more could be done to elaborate on the poverty implications in Stage I economic impact epIAs, even if it is accepted that estimating deeper poverty impacts in Stage II epIAs is much more of a challenge that requires considerable skill, thought, and planning.

We hope that these 'good practice' guidelines will help economize on the scarce resources available for carrying out impact assessments. To make it easier for readers to distil lessons from the text, a capsule summary of each of the 'good practices' discussed is given at the end of each section. This is supplemented by a list of key references that the reader can consult for further insights. CGIAR has recently introduced a performance measurement system, which includes components to document the extent and quality of epIAs and how well an impact assessment culture is being developed. Alongside this development SPIA hopes that the guidelines will assist centers to improve their performance.

The guidelines suggest that research institutes should plan to invest up to 2.5–3% of their resources on impact assessments, with epIA receiving about half of this and the balance going for *ex ante* impact assessments and priority assessments. Effective epIAs are done best with dedicated core resources and staff (rather than *ad hoc* project resources and staff) and they should be located either in a socioeconomics program or in an impact assessment unit.

Placing epIA in the research domain of research centers will help to engender an impact culture because scientists will then see epIA as a scientific exercise and not as a subjective undertaking that mainly marshals information for press releases or as a component of compliance auditing.

These guidelines indicate that there are many remaining methodological challenges in epIA that require further research. The most acute need is to improve the quantity and quality of epIAs in certain research areas, including NRM, livestock, post-harvest, policy, and capacity building, which have had limited epIA relative to the large amounts spent on these important subjects. The linkages between epIA and priority setting in general, and between epIA and ex ante impact assessment in particular, also need strengthening.

SPIA is grateful to the authors, and especially to Tom Walker who led the process, for producing an excellent document. We also thank the centers for sharing their experiences with epIA and for commenting on drafts of these guidelines.

Jim Ryan

Chair
CGIAR Science Council
Standing Panel on Impact Assessment

1. Introducing epIA

Defining the *ex post* impact assessment of agricultural research

Ex post Impact Assessment (epIA) is a specialized area of evaluation designed to identify and measure the consequences resulting from a program or project's earlier interventions. The defining characteristic of epIAs is their timing (Boardman et al., 2001) as epIAs take place after a program or project has generated the intervention being assessed and sufficient time has elapsed and experience accumulated to assess the intervention's performance in terms of longer-term economic, social, and environmental consequences. EpIAs contribute primarily to accountability by demonstrating impact to donors and other stakeholders, and secondarily to learning about the effectiveness of agricultural research. The impacts of an intervention may be positive or negative, primary or secondary, direct or indirect, and intended or unintended (OECD, 2002).

Public-sector national and international research that contributes to agricultural development has a long and distinguished record of epIA (Dalyrmple, 1975; Anderson et al., 1988; Pingali, 2001). As with any long-standing tradition based on impressive accomplishments, there is always room for improvement, and epIAs of public-sector research in developing country agriculture are no exception. This document aims to formulate a set of principles and strategic guidelines for the epIA of agricultural research of the type conducted by the international agricultural research centers (IARCs) funded by the Consultative Group on International Agricultural Research (CGIAR). This mandate encompasses research on biodiversity, crops, fish, forestry, livestock, natural resources, and policy. The setting is interdisciplinary research conducted by biological, physical, and social scientists who aim to achieve developmental impacts mainly for alleviating poverty, enhancing food security, and for promoting environmental sustainability.

One of the motivations for doing epIAs on publicly-funded agricultural research systems is to meet donors' demands for more direct evidence that their large investments in agricultural research are positively affecting the Millennium Development Goals (MDGs) related to poverty alleviation, food security, environmental sustainability, improving health, and generating employment (Raitzer and Winkel, 2005). However, the donors' preference for documenting consequences further along the impact pathway poses methodological challenges for practitioners in attributing such consequences to selected technology or policy interventions. Although these guidelines do not provide any silver bullets to resolve the challenges confronting impact practitioners in carrying out epIAs on agricultural research, they do point to a way of doing business that cost-effectively addresses these challenges.

About this guidance document

The purpose of this guidance document is to help improve the usefulness, the quality, and the quantity of epIA in public-sector agricultural research in developing countries. As such, these guidelines are written primarily for impact practitioners in research organizations that engage in agricultural research of the type conducted by the CGIAR (or simply, CG) centers. Although the discussion and examples cited here mainly address CG-center work, the same principles of good practice are widely applicable to all forms of public-sector agricultural research in developing countries. Institutionally, these include the IARCS, national agricultural research systems (NARS), regional agricultural research organizations, and the networks in developing countries that are involved in agricultural research with the goal of generating development impacts.

This document dispenses only general advice on the conduct of epIAs and so is no substitute for academic texts on the principles of research evaluation (Alston et al., 1998). Nor do the guidelines provide detailed instruction as provided in practitioners' manuals such as Masters et al. (1996).

The purpose of these guidelines is not to supply a detailed step-wise 'how to' manual for carrying out epIAs; but rather to attempt to discuss the 'what' and the 'why'. The emphasis is on key conceptual and strategic methodological issues associated with carrying out epIAs of research interventions in agriculture. Thus, research managers, decision makers, other stakeholders, and donors to public-sector agricultural research in developing countries should find these guidelines informative and useful.

Given its focus on strategic issues related to epIA, these guidelines complement recent and forthcoming work at the Australian Centre for International Agricultural Research (ACIAR), the International Maize and Wheat Improvement Centre (CIMMYT), and the International Potato Center (CIP) that give operational guidelines for the institutional conduct of epIA (ACIAR – Gordon and Davis, 2007), that describe best practice in the form of field guidelines (CIMMYT – La Rovere and Dixon, 2007), and that takes the form of a structured manual on economic rate of return assessments (CIP – Walker et al., forthcoming).

This guidance document is intended to be normative in the sense that it provides advice for planning, conducting, and institutionalizing epIA. These guidelines should not be interpreted as being about imposing 'best practices', as this would be a barrier to innovation; but rather as principles of 'good practice' which, if met, will enhance the quality and practice of epIA in agricultural research. Principles of good practice under each theme/topic are therefore highlighted and, where appropriate, supported with examples of studies as markers towards the development of good practice in epIA. In places, pertinent details are given in boxes and annexes. Succinct statements that summarize the discussion are presented at the end of each section along with key references.

Limitations

This document has several limitations. It is a large undertaking to try and develop epIA guidelines that cover all the possible focal and thematic areas of agricultural research. The two main types of development-oriented agricultural research (as conducted by CG-centers) are technology-oriented research that leads to tangible outputs in the form of improved products, practices, or methods, and policy-oriented research which leads to improved knowledge, information, and recommendations. Policy-oriented research receives less attention in these guidelines than technology-oriented research, although elements of these guidelines can be applied to policy research.

The guidelines are based on observations in the general literature that the lack of comprehensive thematic coverage on the epIA of agricultural research is a chronic problem (Alston et al., 2000). Some of the limitations of these guidelines are a reflection of this lack of coverage with, for example, the majority of epIAs found in the literature focusing on the genetic improvement of major field crops. Relative to the level of investment, the biological control of exotic pests is also well represented. There have been fewer epIAs on crop and integrated pest management (IPM) research, although these subjects do have a profile in the literature. However, only a few epIAs have been carried out in the substantive areas of livestock research, natural resource management (NRM) research, post-harvest research, and policy research. For some of these underrepresented areas, such as NRM, research on the methods for measuring impact appears to be far more abundant than the number of epIAs actually carried out (Shiferaw et al., 2005), although SPIA has recently completed a series of case studies on the impact of NRM research (Waibel and Zilberman, 2007). In contrast, robust methods are lacking for carrying out epIAs on policy-oriented research (SPIA, 2006a). As methods advance in some of these other areas and as experiences are documented, there will be an opportunity to update this guidance document to derive and promote principles of good practice in epIA to broaden and expand the thematic coverage.

Agricultural research generates public goods other than technologies and policies. Foremost among these research-derived public goods are building up the capacity of agricultural research organizations and germplasm conservation. This document does not include guidance on carrying out epIAs on these other important components. Further information on these omissions can be found in a number of the key references given at the end of this introductory chapter.

Organization of this document

The guidelines that follow this Introduction are organized under three major chapters, which describe a strategic approach to epIA, the carrying out of epIAs, and the institutional aspects of epIAs. However, readers need not follow this sequence to use these guidelines. The varying interests and roles of impact assessment practitioners, managers, and decision makers mean that some chapters will be more pertinent to one audience group than another. Chapter 2, which discusses and analyses several strategic aspects of epIA, may be of more interest to research managers and donors. Upper-level managers and decision makers may find the institutionalization of epIA in Chapter 4 more useful than the conduct of good practice epIA discussed in Chapter 3, which is mainly targeted at practitioners.

Key references[1]

Alston J.M., Norton G.W., and Pardey P.G. 1998. Science under Scarcity: Principles and Practice for Agricultural Research Evaluation and Priority Setting. • A most comprehensive treatment on agricultural research evaluation for economists.

Alston J.M., Chan-Kang C., Marra M.C., Pardey P.G., and Wyatt T.J. 2000. A Meta-analysis of Rates of Return to Agricultural R&D, Ex Pede Herculem? • One of the many strengths of this meta-analysis is its evaluation of the determinants of the variation in rates of return to investing in agricultural research.

Anderson J.R. Herdt R.W., and Scobie G.M. 1988. Science and Food. The CGIAR and its Partners • The first comprehensive impact assessment of the CGIAR, 20 years after its inception in 1968.

Dalrymple D.G. 1975. Measuring the Green Revolution: The Impact of Research on Wheat and Rice Production. • One of the first and most insightful epIAs ever conducted on the Green Revolution by a donor representative to the CGIAR.

Evenson R.E. 2001. Economic Impacts of Agricultural Research and Extension. • A comprehensive meta-analysis on the economic impact of agricultural research.

Gordon J. and Chadwick K. 2006. Capacity Building Evaluation.• A rare quantitative assessment of an area that resists facile analysis.

Gordon J. and Davis J. 2007. ACIAR Impact Assessment: Guidelines for Practitioners. (Australian Centre for International Agricultural Research) • Guidelines for epIA at a bilateral donor research organization, which arguably has conducted more impact assessments on agricultural research than any other organization.

Griliches Z. 1958. Research costs and social returns: Hybrid corn and related innovations. • The pioneering application in technology-oriented epIA that focused on hybridization, a new method of invention.

Koo B., Pardey P.G., Wright B.D., and others. 2005. Saving Seeds: The Economics of Conserving Crop Genetic Resources *Ex situ* in the Future Harvest Centers of the CGIAR. • Pioneering analysis of relevant issues in a challenging area.

Ravallion M. 2001. The mystery of the vanishing benefits: An introduction to impact evaluation. • Not related to agricultural research, but a valuable teaching aid that emphasizes the role of different disciplines in contributing to impact assessment.

1 Full references are found in the Bibliography section at the end of these guidelines.

2. Describing EpIA and Rationalizing a Strategic Approach

This chapter is divided into a lengthy, but not exhaustive, description of (epIA); and a rationale for choosing a strategic approach to epIA for international or national research institutes working in the public sector on agriculture, livestock, fisheries, or forestry in developing countries.

The description starts with a typology which forms the basis for these guidelines and then continues by characterizing the outcomes and consequences along an impact pathway by type of epIA. The description of epIA then turns to general methodological issues and concludes with a discussion of the attributes of and standards for good practice whilst carrying out epIA. The second part, on choosing an approach to epIA, begins with a discussion of the pros and cons of impact accounting with individual success stories focusing on specific technologies and policies. It ends with a presentation on the desirability and feasibility of carrying out epIAs on programs. The sequencing of epIAs in a structured fashion is a theme that runs through both parts and several of the sections.

A typology of *ex post* impact assessments in agricultural research

In the literature on the impact of agricultural research, epIAs are represented by several types of analyses that range from disaggregate research-derived, technology- and policy-oriented assessments to aggregate econometric evaluations of the effects of research expenditures on productivity. At the risk of oversimplifying the issue, a typology of epIAs related to agricultural research is presented in Table 2.1 giving four genres of studies. This typology not only establishes a basis for subsequent discussion but also conveys information on what is emphasized and what is viewed as desirable in these strategic guidelines.

Aggregate economic rate of return assessments

A popular method amongst economists for carrying out epIAs, this takes a statistical approach that relates productivity changes to earlier investments in agricultural research (Alston et al., 2000; Evenson, 2001). Most such studies are carried out at the level of a commodity, a group of commodities, or for the agricultural sector as a whole. Typically, aggregate economic rate of return studies do not draw on experimental information for evaluating benefits, do not refer to the adoption of technologies attributed to the research, and do not feature interdisciplinary research as their elaboration requires specialized skills in statistical methods – particularly time series data analyses (Alston et al., 1998). Reliable secondary data on agricultural output, research expenditure, and other determinants of productivity change are essential to support these applications. Economic rate of return estimates based on aggregate assessments tend to be significantly higher than estimates with a disaggregated (specific

Table 2.1. A typology of ex post impact assessments in agricultural research.

Level of the assessment	Primary objective is to document: Productivity and profitability	Selective high order impacts
Macro: Generic technologies/ policies or agricultural research as a whole	Aggregate economic rate of return	Aggregate multi-dimensional impact
Micro: One or more well-defined technologies or policies	Disaggregate economic rate of return	Disaggregate multi-dimensional impact

technology) focus based on cost–benefit analysis (Alston et al., 2000).

There is a time and place for aggregate economic rate of return assessments. However, although in principle such assessments complement the other three forms of epIA described in Table 2.1, they do not come within the purview of these guidelines for two main reasons. Firstly, because of their lack of technological or policy specificity these assessments are not a viable option for documenting impact for accountability purposes or for enhancing learning for technology- and policy-oriented research at the programmatic level of research institutes. Secondly, national or regional data on production and inputs in developing countries are often not sufficiently reliable to support this approach.

Aggregate multi-dimensional impact assessments

Aggregate multi-dimensional impact assessments are rarer than aggregate economic rate of return studies, but they share the common characteristic of relying almost exclusively on secondary regional- and national-level datasets to examine non-economic rate of return consequences associated with or caused by technological change. The term 'multi' is used loosely here as some of these assessments are highly selective and focus on one or, at most, a cluster of related consequences at a macro level (Gollin, 2006). The economic rate of return to investing in agricultural research may also be estimated in these studies, but such estimates are often a means to estimating other effects.

Assessments in this genre include those by Fan et al. (2001), Fan (2002), and Thirtle et al. (2003) who investigated the nexus between agricultural research expenditures, farm-level productivity, and rural and urban poverty. Such assessments are increasingly topical because several of the MDGs may only be quantifiable (or most easily quantified) at this macro-level with regard to changes in the level of investment in agricultural research (Rosegrant et al., 2007). In spite of their desirability, aggregate multi-dimensional impact assessments are referred to, but are not discussed explicitly, in these guidelines for the same reasons that aggregate economic rate of return assessments receive little attention.

Disaggregate economic rate of return assessments

Disaggregate economic rate of return assessments focus on one or more well-identified technologies or policies that have been generated by investments in agricultural research. The method of analysis is project appraisal patterned after Griliches (1958). Typically, researchers engaged in disaggregate economic rate of return studies:

- Draw on multiple data sources for evaluating benefits
- Invest in primary data collection to generate estimates of adoption and/or benefits
- Can interact with scientists who are involved in developing the research product.

Thus, such assessments have a potential for both accountability and for learning, particularly about the design of the assessed technology or policy. As the emphasis is on documenting the economic rate of return in agricultural research, these assessments represent an additional step beyond studies on the adoption of specific interventions. In general, disaggregate economic rate of return assessments generate information on the size of direct impact resulting from research-induced changes in impact indicators and on partitioning impacts between those on consumers and those on producers. Aside from indirect price effects, they usually do not address deeper non-economic consequences along the impact pathway.

The conduct of disaggregate economic rate of return assessments is one of the staple tasks of practitioners who carry out epIAs on agricultural research. The majority of studies included in recent reviews of the impact of CGIAR-related agricultural research belong to this genre (Maredia and Raitzer, 2006; Raitzer and Kelley, 2008). In spite of their frequency in the literature, such assessments are still needed 'to provide a bottom line' as there are few if any substitutes to inform on the productivity of investing in agricultural research over time, space, and research area in the context of developing countries where data are often sparse and incomplete. Of the four types of epIA described in Table 2.1, disaggregate economic rate of return studies command the lion's share of attention in these strategic guidelines. Maintaining and improving the quality of these assessments is emphasized, especially in Chapter 3. Making disaggregate economic rate of return assessments more like (or at least a more informative precursor to) disaggregate multi-dimensional impact assessments is another theme that weaves its way through these guidelines.

Disaggregate multi-dimensional impact assessments

Disaggregate multi-dimensional impact assessments are rarer than disaggregate economic rate of return assessments. These multi-dimensional impact assessments embrace a diverse set of inquiries exemplified by the effects of technological change on growth linkages (Hazell and Ramasamy, 1991), labor markets and migration (David and Otsuka, 1994), regional producer welfare (Renkow, 1994), and poverty (Adato and Meinzen-Dick, 2007). Such research may take several years and, almost always, builds on field work and studies conducted as much as 20 years ago (Bourdillon et al., 2007). Some impacts, such as effects on poverty, nutrition, health, and the environment, are notoriously difficult to document and attribute. For this reason, disaggregated multi-dimensional impact assessments focus mainly on 'larger' or more visible technologies, such as Green Revolution technological change, that are perceived to offer better prospects for the documentation of impact.

The use of mixed methods is a distinguishing trait of multi-dimensional technology impact assessments. Another defining characteristic is a broader capacity for learning with expected results potentially relevant to technology transfer, policy, and other areas related to economic development. In contrast, learning in disaggregate economic rate of return inquiries is linked to accountability for the resources invested and confined to key strategic issues, especially technology design and the allocation of research resources.

In terms of both budgetary support and human capital, a disaggregated multi-dimensional impact study can be quite demanding and costly compared to a disaggregated economic rate of return assessment. For example, each of the five micro-studies on the impact of agricultural research on poverty described in Adato and Meinzen-Dick (2007) is reported to have cost US$200,000. The supply of these studies is more likely to be constrained by lack of funding than the other types of epIA given in Table 2.1. Although they do not figure as prominently in these strategic guidelines as disaggregated economic rate of return studies, disaggregate multi-dimensional impact studies do receive considerable attention.

Stage I and Stage II epIA

In addition to the typology presented in Figure 2.1, this guidance document distinguishes between Stage I and Stage II types of epIA to highlight the taking of a sequential approach to epIA. These two types differ in scale, scope, complexity, and timing of impact assessment.

- Stage I epIAs are technology-focused studies that assess intermediate impacts at some geographically aggregate scale (district, region, country) after the dissemination and adoption have occurred and where there is evidence of costs and benefits (economic, social, environmental) at the adopter-level. The disaggregated economic rate of return studies described above fall under this category[2].
- Stage II epIAs are studies conducted to assess the 'big picture' community-level impacts along the impact pathway when adoption is sufficiently scaled up to see changes in mission-level goals (i.e., poverty, environmental sustainability, health) as reflected in macro-level indicators. The disaggregated multi-dimensional impact assessments described above fall under this category. The distinguishing features of Stage I and Stage II are further described in the second section of this chapter in the context of the 'impact pathway'.

Typology and terminology

Unless stated otherwise, 'economic rate of return' and 'multi-dimensional impact

2 But Stage I epIAs can be multi-dimensional and need not focus exclusively on economic impacts (see for example Dey et al., 2007).

assessments' refer to the disaggregate evaluations which occupy center stage in these guidelines. Stage I and Stage II epIAs highlight the sequential approach to epIA – the distance along the impact pathway.

The term 'research interventions' is used here to refer to technology outputs (i.e., the outputs or consequences of research embedded in material objects, products and practices), and policy outputs (i.e., outputs or consequences of research that lead to improved understanding and an enhanced body of knowledge to guide decisions and achieve a desired outcome) selected for epIA. Such interventions are also called 'selected technologies' and 'selected policies'. Selected technologies, selected policies, and research interventions are used interchangeably throughout these guidelines.

Documenting consequences along the impact pathway

The ultimate goal of research undertaken by national and international agricultural research agencies is to improve the living standards of the poor and the environment in which they live. Figure 2.1 gives a conceptual picture of how a development-oriented, global public goods type research intervention leads sequentially to the generation and then uptake of research results and impacts on ultimate goals. This diagram gives a simplified representation of the complex interrelationships of the many factors involved in creating an impact.

The tapering thickness of the arrow spanning this two-dimensional time and impact pathway reflects the declining influence of a given research intervention on the subsequent stages and increasing complexity and difficulty of attributing an impact to a specific intervention. In reality, it is usually not possible to determine a scientifically sound, discontinuity-free, cause-and-effect relationship between a research intervention and the changes observed at an aggregate level. This measurement problem is referred to as the 'attribution gap' in the literature. Baur et al. (2001) and EIARD (2003) maintain that a central task of impact assessment is to establish highly plausible links between a research effort and the observed changes along the impact pathway (in both Stage I and Stage II epIAs). However, as illustrated by some recent studies, a principle of good practice epIA is to enhance the rigor of establishing the links along the 'input–output–

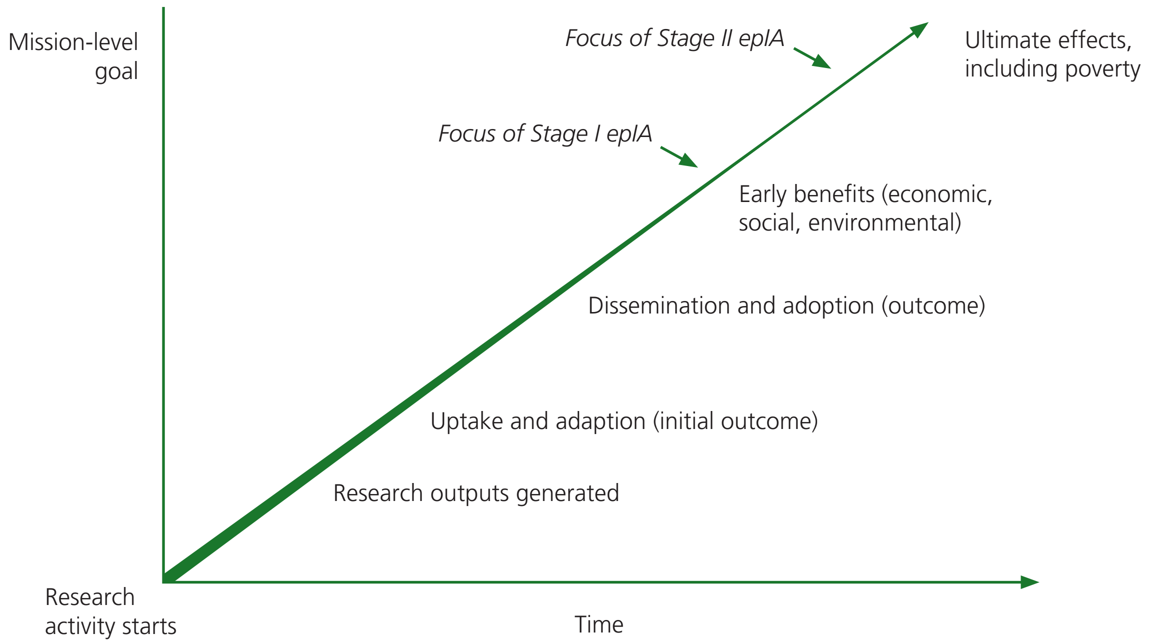

Figure 2.1. Impact pathway for development-oriented global public good agricultural research.

outcome–impact' pathway beyond 'plausibility' to 'substantially demonstrated' impacts (Maredia and Raitzer, 2006; Raitzer and Kelley, 2008).

Examples of impact pathways related to CGIAR research

The CG-centers update their Medium Term Plans (MTPs) each year and these serve as important components for monitoring and evaluation (CGIAR 2006a). All projects included in these plans provide a brief description of the most plausible impact pathway, from problem identification through to the intended ultimate goals. These *ex ante* (predicted) impact pathway descriptions are specific to the research proposed and provide a description of the planned outputs, intended users, expected outcomes, and likely impacts.

Conditioning factors, such as policy and institutional constraints, or enablers that underlie the probability of achieving desired outcomes and impact, also need to be identified. Ideally, the same project description should specify the target ecoregion(s), the intermediate and end-user beneficiaries, and describe the center's potential role vis-à-vis partners and other actors to help ensure outcomes and impact (CGIAR, 2006b). Many of these same considerations and narrative descriptions – based on evidence – should also feature in good practice epIAs.

The pathway from research to outputs to outcomes to impact is a cumulative result of 'cause-and-effect' relationships between many players and factors as depicted in Figure 2.2 for the five categories of CGIAR undertakings. These five categories are: enhancing NARS, research on sustainable production, germplasm improvement, germplasm collection, and policy research. The task of epIAs is to map out, based on evidence, the links between a specific intervention and the relevant outputs, outcomes, and realized impacts, by identifying the cause-and-effect relationships between these different factors.

Figure 2.2 needs to be clarified through making a number of observations:

- Firstly, the research-to-impact pathway involves several intermediate steps, and although not depicted in the figure, an intervention can potentially have multiple outputs, outcomes, and impacts. These can be intended or unintended, and positive or negative. In other words, there can be multiple pathways from research inputs to 'big picture' impacts.
- Secondly, the time lags between research and intermediate steps increase along the pathway. Similarly, the realization of the 'big picture' impacts may take several years, even decades, and is often a result of cumulative effects of many inter-related interventions, outputs, and outcomes, not all of which can be orchestrated by the research institution or predicted accurately. Depending on the time when an epIA is conducted and the complexities involved, an epIA may focus only on Stage I or Stage II impacts. Seldom will any single epIA be comprehensive enough to address all the Stage I and Stage II impacts.
- Thirdly, not all the players and factors listed in this stylized map will be part of all the impact pathways. The 'pathway' by which a research project influences the level of achievement of ultimate goals, i.e., the size of the impact, will depend on the specific outputs generated, the presence (or absence) of uptake of those outputs by immediate users, the magnitude of changes observed at the adopter-level, and the scale of adoption of outputs by end-users. Overall, this generic 'map' depicts the complexity of interrelationships of the many factors involved in creating an impact.

To illustrate these interactions in a less abstract setting, Figure 2.3 presents the impact pathway for the introduction of new soybean varieties in Nigeria (Sanginga et al., 1999) whilst Figure 2.4 gives an example of the impact pathway for policy-oriented research in the Philippines (Templeton and Jamora, 2007). In the example of new varieties (Figure 2.3), the overarching goal of the project was to improve the wellbeing of poor farming families by developing and encouraging the adoption of early maturing high-yielding soybean cultivars. The diffusion of these varieties went hand in hand with the expansion of soybean as a new food and cash crop. Because of the strength of soybeans' rapid diffusion in well-defined regions, it was possible to assess several of the wider

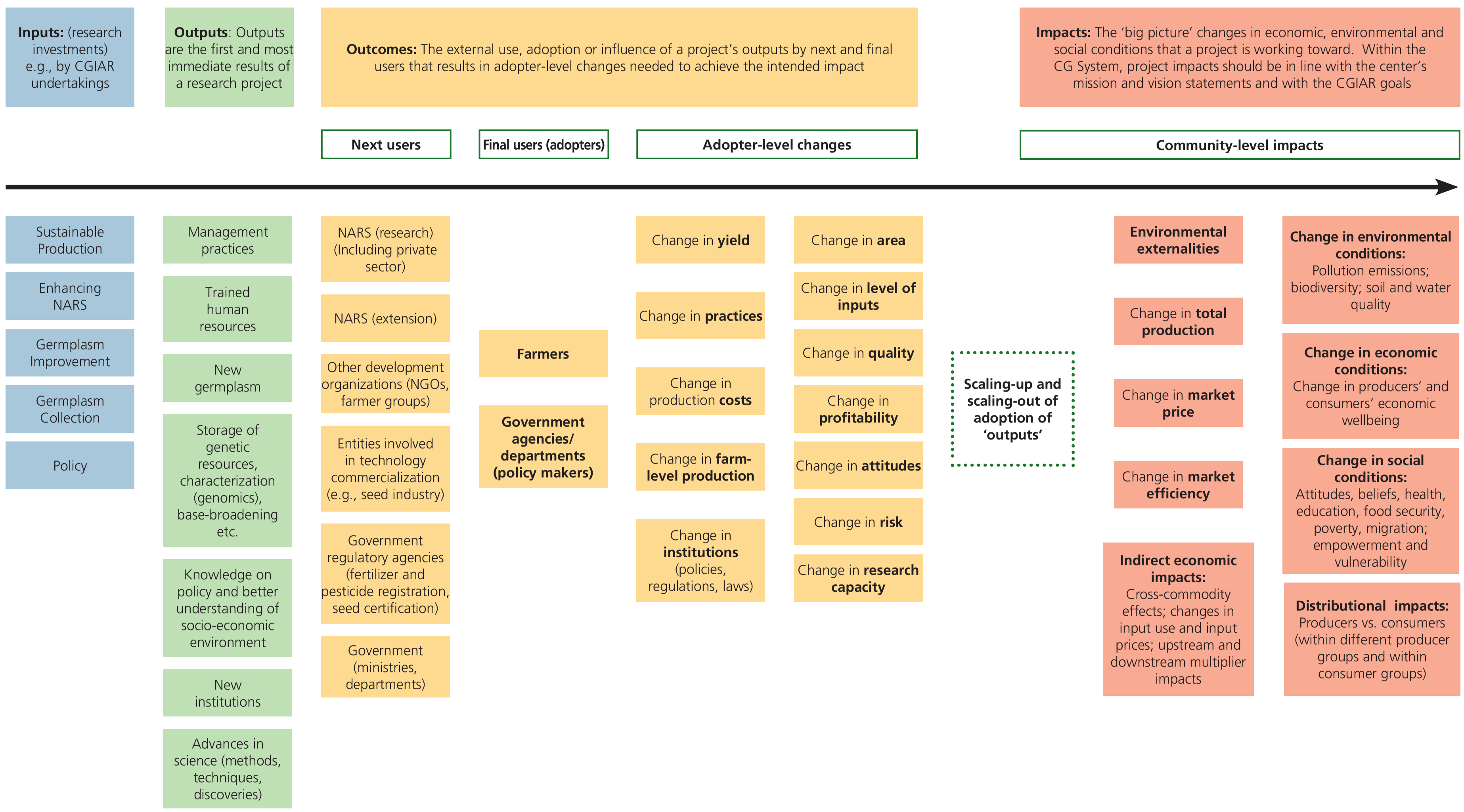

Figure 2.2. A stylized map of factors and players involved along the pathway from research inputs to outputs to outcomes to impact for the CGIAR research portfolio from an *ex post* impact assessment perspective.

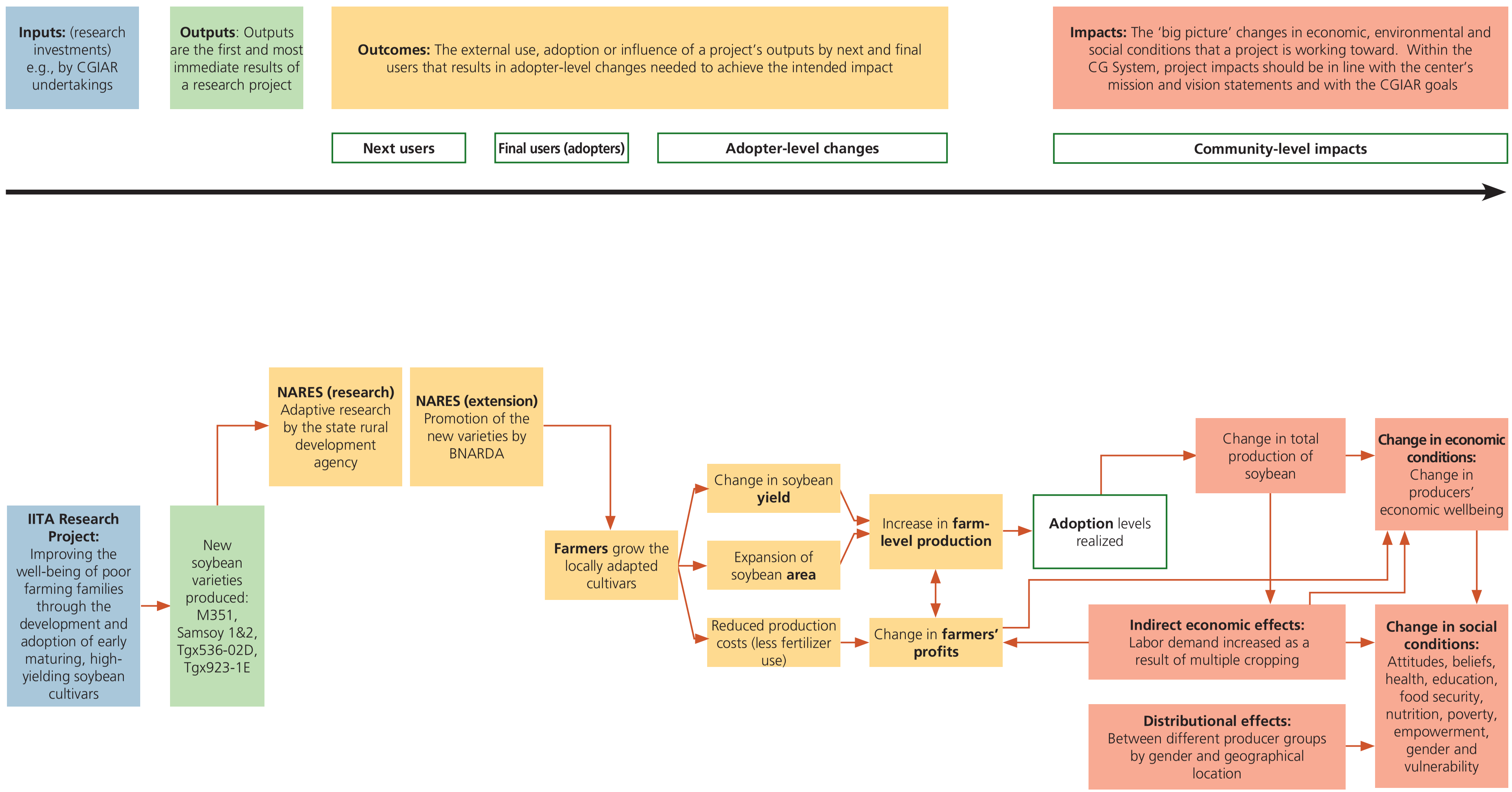

Figure 2.3. Example of impact pathways for a research project focused on developing high-yielding varieties. Source: Sanginga et al. (1999).

Note: IITA = International Institute of Tropical Agriculture; NARES = national agricultural research and extension system; BNARDA = Benue Agricultural and Rural Development Project

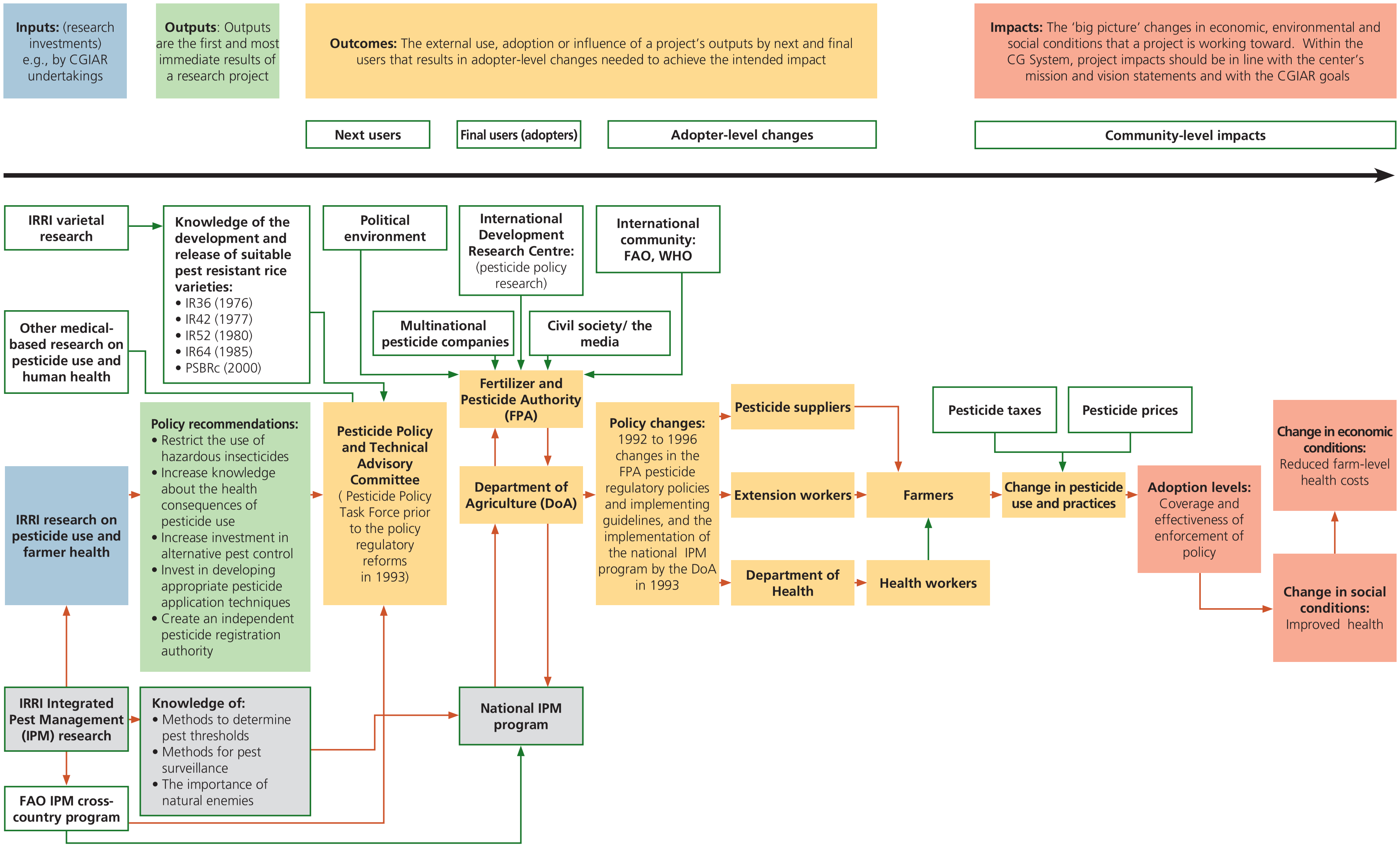

Figure 2.4. Example of an impact pathway realized from policy-oriented research conducted by the International Rice Research Institute (IRRI). Source: Templeton and Jamora (2007.

Note: IRRI = International Rice Research Institute; FAO = Food and Agriculture Organization of the United Nations; WHO = World Health Organization impacts mapped in Figure 2.3 (though not necessarily at a macro-level). One of the challenging areas for this assessment was separating out the contribution of the improved varieties – which is difficult to measure – from the effects of the expansion of soybean growing, which generates more visible consequences.

The example of the impact pathway of policy change in the Philippines traces the pathway from the outputs of the International Rice Research Institute's (IRRI's) research on pesticide use, farmer health, and IPM along the steps that lead to the outcomes (change in regulatory policy and establishment of the integrated pest management national program) and adopter-level and community-level impacts (Figure 2.4). This example also illustrates the role of outputs and outcomes of other types of research (e.g., medical research) and players (e.g., FAO) in influencing the outcomes that resulted in impacts being assessed. Such explicit recognition of the role of other possible research and other types of interventions in realizing the outcomes and impacts gives useful insights for sorting out the issue of attribution.

In contrast to the generic map in Figure 2.2, the specific contexts in Figures 2.3 and 2.4 typically result in a smaller and more incisive set of outputs, outcomes, and impacts. For epIA practitioners, mapping the specific context pays dividends for identifying the most important outputs, outcomes, and impacts on which to focus assessments. These examples also illustrate the varying type, scale, and scope of impacts assessed by epIA studies.

A secondary output of most research projects is the increased capacity of collaborating scientists to use scientific tools and techniques and of extension workers to design and implement effective communication and dissemination activities. Advances in scientific knowledge can be another, if unintended, output. However, as the aim of epIAs is to establish the 'substantially demonstrated' plausibility of a project's impact on the final goal – to determine if poor farming families become better off as a direct or indirect result of a project – in most cases, quantification of the capacity-building and scientific impacts may be beyond the scope of a single impact study. The focus of most technology-oriented epIAs is to quantify, as much as possible, community-level impacts, particularly changes in the economic conditions that can be attributed to a project. While some of the research-induced positive or negative changes in environmental and social conditions may not be quantified or indeed be quantifiable, a brief qualitative assessment can be undertaken. At a minimum, any marked contributions to science and capacity building should be described and catalogued for future assessments that may focus more on these two areas in the spirit of Gordon and Chadwick (2006).

Examples of good practice

Given the inherent complexities involved, developing an impact pathway provides a pragmatic strategy for documenting hard-to-track consequences. Researchers who investigate impact have an abstract, conceptual picture in their mind of what impact is and how it works – a 'map' or hypothesis of how a specific project is linked with various factors and observed changes in target impact indicators. As such, developing an impact pathway can be an important prelude to *ex post* evaluation as it provides a guide to the major focal points of the analysis and to data needs and sources. Therefore, even if not required, undertaking an epIA within an impact pathway framework helps to increase the likelihood that all intended and unintended, positive and negative impacts are identified, and where possible quantified.

Some tools for developing an impact pathway framework in the context of project monitoring and evaluation are briefly described in the subsection below on 'Outcome mapping and participatory impact pathway analysis' (PIPA) (see page 16). For example, participatory impact pathways analysis (PIPA) – a version of outcome mapping – has project implementers and stakeholders working together to agree and define impact pathways, which are then made explicit as impact hypotheses. This type of analysis can also be carried out at the end of a project to reconstruct the actual impact pathways. As this is a relatively new approach, which needs to be further assessed, we are not yet in a position to assert that it represents good practice for epIA. It should be emphasized, however, that either explicitly or implicitly all epIA studies must construct an impact pathway from the research output to impact indicator of interest, including addressing the essential counterfactual[3] and attribution dimensions.

Turning to substantive results, aggregate and disaggregate multi-dimensional impact assessments have done a reasonably good job of documenting the impacts of Green Revolution technological change on aspects of deeper-seated consequences, such as consumption, nutrition, smallholder income, employment, risk reduction, and ecological sustainability (Lipton, 2007). Unfortunately for impact assessment practitioners, most technologies do not resonate across society as markedly as new cultivar ideotypes, such as the high-yielding wheat and rice varieties in Asia, or new field crops, such as soybeans in Nigeria. For technology-oriented economic rate of return assessments, the only impacts that are reasonably transparent for estimating benefits relate to changes in the economic conditions of producers and consumers (see right-hand column of Figure 2.2).

For policy-oriented and capacity-building epIAs, even just elucidating these effects is a challenging task. The problem with documenting outcomes along the impact pathway is related to the incremental and cumulative nature of technological change. It is rare that a technology can be conclusively identified to help numerous beneficiaries cross a poverty line or to make a quantum leap in nutritional and health status. Almost all successful technologies in developing country agriculture contribute to the MDGs, but disentangling and measuring their ultimate effect on each goal is often challenging.

A pragmatic strategy for tackling the seeming intractability of documenting consequences along the impact pathway calls for a sequential research approach to epIA, especially for assessing the impact of technologies. Working in this way, in the Stage I assessment, the direct benefits to users are documented and priorities for assessing other consequences along the impact pathway are identified for later analysis. It must be stressed here that it is a risky proposition to attempt direct assessment of the Stage II outcomes further along the pathway without first executing the Stage I assessment. Initially, it is better to concentrate efforts on demonstrating the direct benefits to beneficiaries unless the technology is expressly designed to respond to one well-defined impact along the impact pathway (Center for Global Development, 2006).

Many research institutes have the capacity to do Stage I epIAs. These usually focus on the benefits accruing to farmers, with extrapolation via simple economic models to the distribution of producer and consumer benefits. It can also be possible to address a few of the Stage II consequences, such as the impact of varietal change on biodiversity. But, in general, the Stage II analysis of deeper effects often requires more specialized technical expertise that may not lend itself to interdisciplinary research, or the mix of disciplines required may be different from that available in the research institute in question. Priorities for Stage II epIA are often episodic and seldom routine. Several distinct technologies may entail different priorities for Stage II studies, even within the same research institute. Lastly, Stage I epIA can be carried out on a shoestring budget (assuming internal evaluation is conducted) with costs in the neighborhood of US$10–30,000 at current prices in recurrent expenses per study (based on 14 epIAs conducted at CIP from the early 1990s to the early 2000s). Doing equivalent quality work in a Stage II study could cost hundreds of thousands of dollars.

The desirability of moving epIA along the impact pathway is unquestioned. As donors want to see ever more comprehensive impact assessments, so ways have to be found to accommodate their wishes even when resources for carrying out these Stage II studies are not forthcoming. One way to enhance the supply of knowledge on deeper consequences can be to encourage (and more widely publicize) research that is

3 The counterfactual dimension or case is discussed in detail on page 39. It is defined there as: what would have happened had the research not been conducted and the resulting technology not existed.

strategically informative about such impacts.

For example, meticulous field research on the effects of pesticides on human health showed conclusively the deleterious effects of heavy pesticide use in rice cropping systems in the Philippines and in intensive potato cultivation in Ecuador (Rola and Pingali, 1993; Crissman et al., 1998). The negative effects of pesticide use on human health were found to be cumulative over time. Technologies that significantly reduce pesticide use could therefore have a large positive impact on human health, but such effects could go undetected whilst documenting impacts over time because of the need to monitor the health of adopters and non-adopters over long periods. However, minimal data such as that on pesticide poisonings can quite easily be monitored, and models based on the earlier work could be used to estimate overall effects. In other words, the previous strategic research could have contributed a building block on which an evaluation of hard-to-document consequences could be based in a later epIA.

Large-scale field efficacy studies on nutrition are another example. A recent study with a multi-year, quasi-experimental design in Mozambique showed that the consumption of orange-fleshed sweet potato by children resulted in significantly improved blood retinol levels (Low et al., 2007). This therefore suggests that the adoption of these kinds of sweet potatoes should not only translate into conventional production or consumer price effects but should also be accompanied by a reduction in Vitamin A deficiency in children and other vulnerable groups. With significant adoption of orange-fleshed sweet potato, the original study could be replicated in the adopting region, although doing this on the same scale would cost upwards of a million dollars. It would be considerably more cost effective to build on the results of the original efficacy study to derive estimates of the nutritional impacts of orange-fleshed sweet potato when the technology is scaled up.

The above are two amongst a number of studies that demonstrate the impacts of field research along the impact pathway from research inputs to impacts on the ultimate beneficiaries (as shown in Figure 2.2). Such studies should be judged on their uniqueness in the literature, on their worth as providing a foundation for Stage II impact assessment when allied technologies are adopted, and on the degree of difficulty of measuring effects.

The above discussion has centered on technology-oriented research. But expanding the umbrella of epIA may also apply to policy research whose value is often derived from blocking or delaying the implementation of 'bad' policy. For example, Benfica et al. (2004) report how well-defined research can play a major role in stopping implementation of new policies that are identified as being likely to hurt poor producers in both the near and medium terms. To come under this umbrella such research would have to demonstrate the deleterious consequences of a specific policy that is about to be implemented. Delaying the implementation of highly distortionary policies can be valuable, as can preventing the removal of welfare-enhancing policies. By the same token, this reasoning does not imply that epIA should be expanded to include all policy-oriented research.

Good practice 2.1. EpIAs of specific technologies should be viewed as a sequential process. In general, Stage I epIAs focusing on the direct effects on beneficiaries should be rigorously carried out before other consequences along the impact pathway are evaluated in Stage II epIAs.

Good practice 2.2. State-of-the-art applied field work that contributes to the elucidation of the social, environmental, and distributional consequences of promising, but as yet unadopted, technologies and policies should be recognized as supporting future epIAs.

Good practice 2.3. While some of the research-induced positive or negative changes along the impact pathway may not be quantified or indeed quantifiable, a brief qualitative assessment should be undertaken or at least alluded to in the analysis and reporting of results.

Key references

David C.C. and Otsuka K. 1994. Modern Rice Technology and Income Distribution in Asia. • A thorough analysis of the effects of the Green Revolution on labor markets and migration.

Hazell P.B.R. and Ramasamy C. 1991. Green Revolution Reconsidered: The Impact of the High-yielding Rice Varieties in South India. • A rigorous confrontation of perceptions with empirical facts from the perspective of social accounting and general equilibrium analysis.

Renkow M. 1994. Technology, production environment, and household income: Assessing the regional impacts of technological change. • A comprehensive regional analysis of the effects of high-yielding varieties of wheat on the distribution of income for producers.

General methodological issues

Methods employed for epIA depend on the type of assessment and its purpose. Cost-benefit analysis is the dominant method used in economic rate of return assessments. In contrast, mixed methods that reflect sound disciplinary practice provide the analytical engines for multi-dimensional impact assessments.

Regardless of the method, epIA practitioners should be aware of two general issues related to conducting epIAs.

- Firstly, carrying out epIA does not mean that practitioners need to make a compelling case for the success of a research intervention. As far as is possible, they should provide credible, objective evidence about the impact of an intervention, which may turn out to be substantially more or less successful than previously thought. The unbiased reporting of insignificant results is just as important as documenting more publishable significant findings. Impact assessment is definitely not about finding only positive impacts – it is about documenting and reporting whatever is found (Ovretveit, 1998). Conversely, going too far along the impact pathway searching for consequences over which the intervention can have only limited leverage could generate insignificant results whose estimation represents an ineffective use of resources.
- Secondly, incomplete information is one of the defining characteristics of epIA. Many assumptions are needed to compensate for informational deficiencies. A handful of these suppositions will often be critical to the results. When assumptions are chosen, they may be optimistic, realistic, or conservative about the settings for key variables.

Common sense suggests that realistic settings should be chosen. But in assessing impact realism is often hard to quantify and so conservative settings are usually a better choice. We want to generate estimates that should understate the size of economic impact: “We do not know how large the impact is but we are fairly certain that it is at least this much.” The serious error to make here, which is a cause for concern, is to grossly overestimate the size of impact. The cost of understating impact does not appear to be that large. However, even conservative assumptions do not guarantee immunity from overestimating impact. But conservatism is one of the few means at practitioners’ disposal to protect against the making of egregious errors.

The following text discusses the carrying out of epIA in the three chronological stages of precursor assessments, Stage I epIA, and Stage II epIA. Outcome mapping and impact pathway analysis, commonly used for planning, monitoring, and evaluation (precursors to epIA) are a means to strengthen the linkages between monitoring and evaluation and epIA. Such analysis mostly applies to shorter-term donor-funded adaptive research projects and its skillful use can lay a sound foundation for subsequent epIAs. Cost–benefit analysis is briefly discussed in the context of economic rate of return assessments (as undertaken in Stage I epIAs). The contributions that multi-disciplinary livelihoods approaches can make for multi-dimensional impact assessment (undertaken in Stage II) completes the presentation in this subsection.

Outcome mapping and participatory impact pathway analysis

Outcome mapping was originally proposed as a methodology for planning, monitoring, and evaluating development programs by the International Development Research Center (IDRC) (Earl et al., 2001). The proponents of outcome mapping are skeptical about the possibility of attributing impact to programs for social change because of the complexity of the development process and the non-unidirectional nature of resulting and associated change. Hence, outcome mapping focuses on 'outcomes', defined as the changes in the behaviors, actions, or relationships of the key stakeholders (known as 'boundary partners') who can be influenced by the program. This definition of 'outcomes' is very similar to the current usage in the CGIAR (CGIAR, 2006b).

A typical application of outcome mapping will see the project team and its partners clarifying the selected project's vision as the desired long-term impacts and identifying key boundary partners and progress markers for tracking performance. This occurs during the project's planning and implementation stages. By monitoring the progress markers and the achievement of outcomes continuously throughout a project's life, outcome mapping both helps the project to achieve its goals and enhances team and program understanding of the intended change process (Salas, 2002). The monitoring of results is a process of reflection, which identifies what is not working (to be left behind), what is going well (to be improved on further), and what is wrong (to be corrected) (Raij, 2004). The continuous collection of data about progress markers during a project's implementation phase potentially provides a launching pad for the *ex post* assessment of outcomes and impacts. For example, in a farmer field school case study in Kenya, outcome mapping was proposed not as an alternative to, but rather to complement a cost–benefit analysis (Nyangaga et al., 2006). Used in this way, outcome mapping can potentially contribute valuable information to confirm attribution in development programs that have as intended consequences social change and an improved understanding of the change process.

PIPA is a recently developed approach to planning, monitoring, and evaluating research-for-development projects that shares some similarities with outcome mapping (Douthwaite et al., 2007a). It involves holding a workshop where project implementers and stakeholders make explicit the anticipated impact pathways. Impact pathways provide hypotheses about cause-and-effect linkages that connect project outputs to outcomes and, under some circumstances, also to impacts (a principal point of difference from outcome mapping). The advantage of using participant-generated impact hypotheses rather than evaluator-generated ones is that project staff and stakeholders are more likely to use and learn from the impact assessment findings. (See http://impactpathways.pbwiki.com for more information on PIPA). Where the quantification of impact is difficult, outcome mapping and PIPA provide interim indicators of the links among outputs and outcomes. This is especially relevant for monitoring and evaluating NRM and policy-oriented research where attributed cause and effect is often arduous and the gestation period to impact is often lengthy.

One disadvantage of using outcome mapping and PIPA is the presumption that scientists know what their findings will be. If this is not the case it is difficult to advance beyond a conventional logical framework. Moreover, the value of these approaches (outcome mapping and PIPA) needs to be further assessed in the context of strategic research programs that are targeted to generate long-term international public goods (where boundary partners are often undefined and it may not be possible to identify all the potential outputs, outcomes, and impacts to be realized from a research activity 10–20 years down the road). More experiences and applications in the use of these planning, monitoring, and evaluation tools need to be reported in the literature so that their contribution to and relevance for epIA (in the specific context of agricultural research aimed at generating long-term international public goods) can be rigorously evaluated.

Good practice 2.4. As a precursor to epIA, outcome mapping and PIPA show promise for monitoring, evaluating, and documenting progress along impact pathways to provide a solid foundation for high quality and plausible epIAs, especially for shorter-term adaptive research projects. They complement and are not a substitute for epIA.

Key reference: Douthwaite B., Schulz S., Olanrewaju A.S., and Ellis-Jones J. 2007.

Impact pathway evaluation of an integrated *Striga hermonthica* control project in Northern Nigeria. • An informative example of the use of impact pathway 'logic' which sets the stage for epIA.

Cost–benefit analysis in economic rate of return assessments

Cost–benefit analysis is the staple (and may be the only) method used in technology-oriented economic rate of return assessments. Policy-oriented impact assessments often use more mixed methods, employing both quantitative and qualitative tools. Quantitative approaches generally use economic models to measure the rates of return to research. Although useful for comparing returns across projects, not all impacts are amenable to quantitative analysis. Qualitative evaluations, usually taking the form of retrospective narratives, are more informative and a necessary ingredient for documenting certain types of research impacts such as policy-oriented research.

Although quantitative cost–benefit analysis is encouraged, a rigorous qualitative epIA can also be conducted from the perspective of cost–benefit analysis when the economic valuation of benefits seems impossible but other empirical indicators of influence or adoption are feasible (see forest management example in Box 2.1). Practitioners using cost-benefit analysis in economic rate of return assessments are also encouraged to deploy other methods to evaluate other outcomes and impacts which they believe would enrich their Stage I assessments.

Cost-effectiveness analyses have many similarities to cost–benefit analyses. The main difference is that in cost-effectiveness analysis, the benefits are not estimated explicitly, although information on costs and common units of output, outcome, or impact measures of closely-related alternatives is needed. The level of adoption is as important in conditioning results in cost-effectiveness analysis as in cost–benefit analysis.

There are occasions where the use of cost-effectiveness analysis is relevant. For example, Stein et al. (2007) used it skillfully to evaluate bio-fortification relative to the alternatives of supplementation and industrial fortification to combat Vitamin A deficiency in children in India.

The factors that condition results in cost-effectiveness analysis are similar to those in cost–benefit analysis. Without a minimum threshold level of adoption, research interventions are unlikely to be cost-effective, which is another way of saying that economic rates of return are low. If micronutrient-rich crop varieties are widely adopted, more applications using cost-effectiveness analysis could be forthcoming in epIAs. When benefits are hard to quantify and when close substitutes are available, cost-effectiveness analysis could also become increasingly common in policy-oriented epIA.

Good practice 2.5. Assigning conservative values to important assumptions is recommended when conducting both economic rate of return and multi-dimensional impact assessments.

Good practice 2.6. Quantitative cost–benefit analysis is the standard for assessing economic rate of return assessments.

Good practice 2.7. EpIA practitioners engaged in economic rate of return assessment should complement their quantitative cost–benefit analysis with qualitative methods capable of shedding light on the institutional attribution, uptake, and influence of the research output under analysis.

Key references

Boardman A.E., Greenberg D.H., Vining A.R., and Weimer D.L. 2001. Cost–benefit Analysis: Concepts and Practice. • A comprehensive yet practical introduction to cost–benefit analysis.

Gittinger J.P. 1982. Economic Analysis of Agricultural Projects. • An old classic, full of practical advice on the use of project appraisal in agriculture.

Stein A.J., Sachdev H.P.S., and Qaim M. 2008. Genetic engineering for the poor: Golden rice and public health in India. • Although not an epIA, this *ex ante* analysis clearly demonstrates the use of cost-effectiveness analysis to assess the impact

Box 2.1. Criteria and indicators for sustainable forest management – an example of qualitative cost–benefit analysis

The work in this study by the Centre for International Forestry Research (CIFOR) is an excellent example of qualitative cost-benefit analysis done using rigorous scientific techniques. Criteria and indicator research has gained wide recognition and its results have been adopted by many organizations. Although in this case the researchers were not able to show the precise impact, they were able to show extensive technology uptake, improved forest policies and management at the national and global levels, and the generation of significant international public goods.

The researcher traced the impact pathways as a means of identifying clients of the research and their needs. The next step was to analyze selected clients in more detail. Key staff in the certification agencies were interviewed to ascertain the role that research had played in developing their standards and what the likely outcome of the certification process would have been in the absence of the research. In addition, project documentation, published documents, and internet resources were reviewed to determine whether Forest Stewardship Council certifiers were actually using the CIFOR tools.

CIFOR's efforts were rewarded. By October 2006, 79 million hectares of forests had been Forest Stewardship Council-certified compared with only 10 million hectares when the study had ended in 1999. More than 84% of these forests had been certified by companies that acknowledge some use of CIFOR's criteria and indicator research in their certification standards or audit processes. More certified forests lie in the temperate and boreal regions of the North, but an estimated 5.8 million hectares are in CIFOR's mandate area of the humid tropics.

CIFOR's research has shown that there can be no universal set of criteria and indicators because forest and socioeconomic conditions vary so greatly. The research instead provided tools and methods to assist national and local criteria and indicator development and adaptation through their Criteria and Indicators Tool Box Series, a comprehensive set of eight manuals in six languages, and accompanying decision-support software to guide users through the process (CIFOR, 1999).

In terms of lessons learned, CIFOR rightly perceived the use of criteria and indicators in regulatory and policy processes to be far more promising for broad 'systematic' impact than promoting the independent adoption of better practices by multiple forest managers.

This case differs from the norm, in which the impact of research is relatively easily assessed where new science-based innovations are clearly defined and where their adoption directly affects production, consumption, and human welfare. In such cases, impact is a function of the number of adopters and the land area covered. Nevertheless, impact may also be achieved indirectly; for example through influencing policies, managerial decision-making, and development interventions. In these situations a relatively small number of adoption events – even a single event – can change the way a system or process works, thus leading to substantial impact.

Source: Spillsbury (2007)

of agricultural research on health in terms of disability-adjusted years (DALYs), an indicator used increasingly to evaluate the health consequences of research interventions.

Livelihood approaches in multi-dimensional impact assessments

Taking a livelihood perspective (see, for example., DFID, 1999; Pasteur, 2001) focuses impact assessment on the effects on people's lives. Improving farmers' livelihoods is not only about increasing income; it also involves better food security, social status and inclusion, asset accumulation, and resilience to sources of vulnerability. A livelihoods approach can be applied to review and evaluate the impact of projects and programs, to see how these affect the livelihoods of the poor, and to determine how impacts can be enhanced. Helping to frame questions to assess the wider context in which technologies work is a key feature of the use of livelihood perspectives for assessing impact. For instance, does a selected project technology change people's ability to cope with shocks or with longer-term trends or changes?

A livelihoods approach can also be used as a checklist of important issues to be considered in designing and conducting an impact assessment, such as:

- Defining indicators
- Measuring changes in the factors that contribute to livelihoods (capital assets, institutional structures and processes, resilience or vulnerability of livelihoods, livelihood strategies and outcomes)
- Understanding how these link to one another and the multiple interactions between factors that affect livelihoods.

Assets are the building blocks of livelihoods. The five widely recognized classes of assets are natural assets (e.g., land, water), social assets (e.g., formal and informal networks), human assets (e.g., education, knowledge, health), physical assets (e.g., equipment, transport), and financial assets (e.g., access to credit, remittances). A person or household needs improvements in some or all of these assets to build a secure livelihood. These 'capitals' interact with the policies, institutions, and processes that have a major influence on the livelihoods of people and the livelihood strategies of households – in other words, the combination of activities that households and their members engage in for living, including agriculture and non-farm activities, and the final livelihood outcomes.

By considering the whole realm of farmers' lives, impact assessment through a livelihoods lens can capture many more of the aspects of impact, some obvious and some not so obvious. For instance, plant breeders may think that their work affects only crop varieties when in fact it directly or indirectly influences several components of rural lives, or 'livelihood assets', not just crop yield (see Mexican example in Box 2.2). The approach taken in Box 2.2 is an example of an early acceptance/user relevance study (project-level).

The use of mixed methods is integral to implementing livelihoods perspectives. Combining quantitative and qualitative methods provides both a quantification of impacts as well as explanations of given outcomes (see Adato and Meinzen-Dick, 2007). Using a variety of methods makes it easier to estimate unintended impacts (positive or negative) as well as intended ones, and to link different levels of analysis and components. Changes in indicators from an innovation (e.g., income, yield) can be assessed in terms of how they contribute to livelihoods, directly (e.g., to income and food) or indirectly (on assets, activities, and ability to cope with shocks).

Livelihood impact indicators are designed to measure changes in households' access to assets, institutional structures, relationships, or livelihoods strategies. These indicators should be:

- Outcome-focused, relating to longer-term targets and addressing the wider context of livelihoods because outcomes are diverse and go beyond simple quantitative changes in variables
- Process based, with on-going progress towards planned outcomes by looking at the quality of the processes that influence the outcomes
- Negotiable and open-ended, looking at negative and positive trends and recognizing context.

A list of commonly-used indicators at different levels and in different contexts is given

Box: 2.2. Operationalizing the livelihoods approach for project-level impact assessment in Oaxaca, Mexico

In the late 1990s, the International Maize and Wheat Improvement Centre (CIMMYT) implemented a project to increase productivity and preserve maize diversity from local traditional landraces in Oaxaca, southern Mexico. The project focused on maize diversity, demonstrations and training, and promoting post-harvest technology (metal silos). In 2006, an impact study applied a livelihood approach to assess any subsequent changes in farmers' welfare and to learn how impact could be enhanced from such projects. A sample of 120 households, semi-purposively and then randomly selected from participants and non-participants, were selected to obtain both with-without and before-after comparisons, and to relate changes in livelihood indicators to an available baseline. Econometrics was used to analyze the effects of the interventions by households' participation in the project and household characteristics.

A clustering technique assembled households into four groups based on similarities in their level of 13 defined livelihood assets. These13 assets were components of the five 'capitals' of local households as described above: e.g., quality of land, water availability and access (natural capital); input use, distance from markets (physical capital); number of family members or education (human capital), links with government programs and networks and derived benefits (social capital), and marketable assets and remittances (financial capital).

The study revealed that a significant proportion of participants were still using the project output (as were some non-participants) as a result of farmer-to-farmer spontaneous diffusion. Information on livelihood changes put the findings in perspective, relating impact with the age of adopting farmers and with a general decline in the importance of maize as a commercial crop, while still retaining its consumption and cultural value. The diffusion of maize seed from the project took place the most in the more remote or less market-connected areas, where there are many poor farmers.

Achieving livelihood impact by local and improved maize germplasm, as assessed through the livelihood approach, was however just one of the goals of the project as it also aimed at increasing knowledge about maize diversity and the use of new participatory research approaches.

Source: La Rovere et al. (2008)

in La Rovere and Dixon (2007) and Adato and Meinzen-Dick (2007).

It is natural to want to quantify the core elements of any framework in impact assessment. One of the challenges facing epIA practitioners who use a livelihoods framework is to document the effects of technological change on the five types of assets (capitals). It is particularly valuable to document the effects of change on the asset holdings of rural households and also to make extrapolations and broader generalizations of results. The livelihoods approach is a useful tool for identifying constraints to adoption and for documenting the multi-dimensional impacts of a specific intervention at the micro-level, such as in villages where an intervention took place. Ethnographies, narratives, and human interest stories often inform about the link between a research intervention and assets, but going further and quantifying these effects at an aggregate level requires specialized statistical skills often combined with experimental designs that cannot be substituted by eclectic methods. As such, more examples of the use of the livelihoods framework in epIA are needed to rigorously define good practice.

Good practice 2.8. Applying a livelihood approach is appropriate where it is important to understand the nature of impacts on poverty at the adopter level and the complex set of factors that determine these impacts on alleviating poverty.

Good practice 2.9. Livelihood approaches provide an overarching framework for analyzing disaggregated multi-dimensional impacts, but epIA practitioners still need to apply validated and accepted disciplinary methods for conducting such analyses.

Key reference: Adato M. and Meinzen-Dick R. (Eds). 2007. Agricultural Research, Livelihoods and Poverty: Studies of economic and social impact in six countries. • An interdisciplinary synthesis of the contributions a sustainable livelihoods approach can make to assessing poverty based on five multi-dimensional impact assessments that report differing successes in their use of the livelihoods framework.

Attributes and standards of good-practice EpIA

This section briefly describes the positive attributes and standards of good-practice epIA. Meta-analyses of epIA in the CGIAR have underscored the importance of transparency and analytical rigor (Maredia and Raitzer, 2006; Raitzer and Kelley, 2008). Analytical rigor does not necessarily imply sophisticated methods. Good data quality and completeness, and clarity in specifying assumptions are crucial dimensions of transparency and analytical quality. In the attributes of good-practice epIA given in Box 2.3, the recurring theme in both transparency and analytical rigor is data, data, and data! EpIAs are selectively data intensive and their quality depends as much on the reliability of the data as on anything else.

Additionally, good-practice epIAs, especially economic rate of return assessments, are defined by:

- A clear description of the components and options that comprise the research-derived technology or policy

Box 2.3. List of the components of transparency and analytical rigor for epIA

Transparency:

- Clearly derived key assumptions – 1) their explicitness, 2) their substantiation
- Comprehensive attribution of data sources by citing 1) adoption data, 2) productivity data, 3) price data, 4) adoption-related cost data
- Full explanation of data treatment by explaining 1) adoption data treatment, 2) productivity data treatment, 3) counterfactual derivation, 4) economic valuation, 5) institutional attribution.

Analytical rigor:

- Representative dataset utilized – 1) its reliability, 2) its comprehensiveness
- Appropriate data treatment – 1) appropriateness of data extrapolation, 2) adequacy of analysis of mitigating factors, 3) adequacy of disaggregation by production environment, 4) adequacy of assessment of adoption related costs
- Plausible counterfactual scenario developed – 1) plausibility of assumptions about substitutable innovations, 2) plausibility of changes due to exogenous causes
- Adequate consideration of mission relevance of benefits – 1) adequacy of analysis of mission relevance of economic benefits
- Plausible institutional attribution – 1) plausibility of institutional attribution.

Source: Raitzer and Linder (2005), pp 25 and 27 as cited in Gordon and Davis (2007)

- The ease of integration into a subsequent meta-analysis
- Generalized causal inference in extrapolating the assessed impact experience to other technologies in the same research area
- Indicative thinking of any distinguishing effects of the technology on consequences further along the impact pathway
- Versatility in informing several groups of users including program scientists, research managers, and stakeholders.

Meeting the requirements set out in Box 2.3 would seem to be a tall order, although examples of good-practice epIAs are readily available. Six can be accessed via the SPIA website (http://impact.cgiar.org) in the form of research briefs. Others are found in the Raitzer (2003) and Maredia and Raitzer (2006) meta-analyses, which feature a rigorous selection of past epIAs. The screened epIAs were evaluated for transparency and analytical rigor. The highest ranking ones were those that 'substantially demonstrated' impact (see Annex A of this paper for a listing). Many of these are Stage I economic rate of return assessments. Recent examples of good-practice Stage I epIAs conducted outside the CGIAR include Mather et al. (2003) and Boys et al. (2007).

In addition to ensuring that an epIA demonstrates the attributes of transparency and analytical rigor, it is vitally important that epIAs are conducted with exceptionally high professional and ethical standards. As one of the tools of evaluation, epIAs must strive to adhere to the following four groups of evaluation guidelines adapted from the International Program Evaluation Standards (Russon, 2001):

1. Utility: to ensure that an epIA serves the needs of intended users and will be owned by stakeholders.
2. Feasibility: to ensure that an epIA is realistic, prudent, diplomatic, and frugal.
3. Propriety standards: to ensure that impact evaluations are conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by the results.
4. Accuracy standards: to ensure that an epIA reveals and conveys technically adequate information about the features that determine the worth or merit of the evaluated program.

Ideally, an epIA will fulfill all these standards. However, in practice, epIAs will not always be able to meet each standard equally. It is thus often appropriate to adapt them to the specific impact assessment situation at hand, which can mean giving lower weights to certain standards while placing more emphasis on others. Whatever, the adaptation of the standards to the specific impact assessment situation should be well considered, openly presented, and explicitly justified.

Good practice 2.10. The hallmarks of good-practice epIA are transparency and analytical rigor. Desirable traits for a good-practice epIA also include the sufficiency of information for meta-analyses; reasoned hypotheses on expected but undocumented consequences; technological generalizability, and readability.

Good practice 2.11. As an evaluation tool, epIA should strive to adhere to the professional and ethical guidelines of utility, feasibility, accuracy and propriety standards.

Key references

Maredia M.K. and Raitzer D.A. 2006. CGIAR and NARS Partner Research in Sub-Saharan Africa: Evidence of Impact to Date. • A meta-analysis, emphasizing transparency and analytical rigor of epIAs on CGIAR-related technological change in sub-Saharan Africa.

Russon C. (Ed.). 2001. The Program Evaluation Standards in International Settings. • A collection of papers highlighting the debate, challenges and issues in applying the International Program Evaluation Standards in international settings.

The rationale for an approach that accumulates the results of success stories

When evaluating government interventions, much of the emphasis on learning about 'what works' rests largely on the fact that 'what does not work' is usually not transparent (not apparent), because bad interventions are rarely documented. In agriculture, technology that is perceived as 'bad' by users does not get implemented because it is not adopted. In terms of epIA, agricultural research is distinctive because its impact hinges on the adoption of research outputs. Research-derived technologies that are characterized by non-adoption are equivalent to government investment projects that are not implemented or social programs in which no one participates. Therefore, epIA on agricultural research is often synonymous with success stories where success equates to adoption.

Stressing success stories as an approach to epIA in agricultural research is based on the uncertain nature of the potentially large benefits from agricultural research. Research is essentially a game of chance with nature, and nature is often niggardly. In a portfolio of agricultural research projects, some will result in technologies that farmers adopt while others may not be successful in generating practical results in the short or medium term (even though they may be successful in contributing towards scientific learning and capacity-building outcomes and impacts).

When we define projects at the level of specific technologies, the distribution of the results of agricultural research is often highly skewed. For example, between 1932 and 1979, 182 improved potato varieties were officially released in the United States and Canada. Only 30 of those varieties were ever popular enough to account for more than 1% of area planted with potatoes in any year from their date of release to 1989 (Walker, 1994). If we took a random sample of these varieties as a reflection of performance, we could easily miss the important ones that have generated benefits. The benefits from the outcomes of agricultural research projects are not normally distributed – they are severely and positively skewed.

The skewed distribution of success is one of the most important known facts in agricultural research. The benefits from only a handful of research areas have more than paid for all the investment by donors in the CGIAR (Raitzer and Kelley, 2008). The benefits from one project has contributed about 80% of the CGIAR's impact on agriculture in sub-Saharan Africa (Maredia and Raitzer, 2006) whilst the net present value (NPV) of one successful project at another center accounted for more than 60% of the value of all other successful epIAs documented in the same center (Fuglie et al., 1999).

Almost any biological technology that is accepted rapidly by a sizable number of farmers will generate sufficient returns to cover not only its own costs but also the costs of other research projects that do not ultimately generate practical results for farmers. In general, documenting key success stories is essential to arrive at an initial understanding of the productivity of an agricultural research program. Although every agricultural research project will not necessarily successfully generate a practical impact, it is reasonable to expect mature agricultural research programs to document at least one success story.

Impact assessment based on success stories is a common practice in the agricultural economics literature, and impact accounting by charging net benefit streams of project success stories to all costs of the program, the institute, and even the system in which they originate, is increasingly practiced (Walker and Crissman, 1996; the U.S. Forest Service as cited in CGIAR 2002; Raitzer and Kelley, 2008). At the risk of building a straw man (a weak argument), the following are the most common critiques of impact accounting based on success stories:

- Won't an emphasis on success stories lead to an overestimation of the profitability of investing in agricultural research?
- Wouldn't it be better to take a random sample of projects to generate an unbiased estimate of economic performance?
- Couldn't we learn more from investing in case studies of technology 'failure'?
- Why should past success be a good predictor of future performance?

The first two questions pertain to accountability and the latter two refer to learning. Critics of the success story approach rightly point out that finding a high economic rate of return of 50% for a specific technology of a program in an agricultural research institute does not show that the economic rate of return to investment in the institute has been 50%. One highly profitable success story cannot be used to extrapolate the measure of success to the entire agricultural research program or institute. But, based on the earlier discussion, taking a random sample of projects may grossly underestimate the impact of an agricultural research institute unless the larger success stories are by chance included in the sample[4].

Ironically, impact accounting with success stories is unlikely to lead to overestimates of programmatic impact. This strategy for programmatic impact assessment should result in a lower bound estimate because not all such impact will be included in documented success stories. The question to ask is: can the profitability documented in the epIA technology cover all the costs of the programmatic research area? If it cannot, then the success story is relatively small, and more technological success needs to be demonstrated in the same program to make a persuasive case that investing in that programmatic area pays. If an attractive economic rate of return, such as 20% or above, can be generated by charging all program costs to the technology, then we have good evidence that investment in the program has paid off. (This assumes that the so-called 'dry holes' do not become 'poisoned wells' that generate losses to society other than the costs of R&D invested in them). The impetus for doing economic rate of return assessments diminishes as success is documented and accountability becomes less of an institutional issue. For such programs, multi-dimensional impact assessments are much more informative than more economic rate of return assessments.

Impact accounting with success stories does not allow the generation of a reliable point estimate of the economic rate of return in investing in research in a particular research area. The approach generates a lower bound estimate of the returns to research.

4 This is because of the skewed distribution of success rate in a research portfolio. If the rate of success was normally distributed across all projects – with a bell-shaped distribution curve – then taking a random sample of projects could be an appropriate strategy.

It is also flexible and can be upgraded over time in a simple spreadsheet format.

An absence of well-documented success stories in a mature programmatic area signals the importance of the need to learn more about why results do not meet expectations in terms of their practical impact. Doing this will focus attention on troubleshooting the allocation of resources for research and management efficiency. For example, at CIP, an absence of success stories stimulated the reevaluation of several research areas described as troublesome from an impact accountability perspective (Chilver et al., 1999; Fuglie et al., 2000; Walker and Fuglie, 2006). The absence of epIAs triggered a broader evaluation. Just because a research area has a public goods character with no alternative means of supply does not justify continuing support if the impact is falling short of expectations.

EpIAs inform about past impact. If past performance in a research area is a guide to the future, then the sparse documentation of *ex post* impacts in the past implies low performance in the future, because interventions that do not work tend not to be documented. Conversely, abundantly documented *ex post* impacts in the past should augur well for future impact because success gets documented. Paradoxically, some donors consider the prospects for future impact equally bright irrespective of the amount of epIA success that has been documented in the past (Raitzer and Winkel, 2005). A well-documented success story may be received with indifference from donors and government officials if they believe that conditions have become substantially different from those when success prevailed. Change is inexorable, but how change could severely alter the impact prospects for different types of research areas is difficult to see unless progress in basic research is accelerating in some areas compared to others.

Studies have shown that scientists tend to be equally optimistic about the prospects for impact in their own research area irrespective of past results. Perhaps donor behavior mimics this optimism. Whatever the case, this behavior is worrying because it suggests that information contributed explicitly by epIA on what works, and conversely by the absence of epIA on what does not work, may be ignored.

Good practice 2.12. An approach that accumulates the benefits of successful research-generated interventions and then 'charges' them to total programmatic costs of both successful and unsuccessful projects is recommended.

Good practice 2.13. The absence of epIA documented success stories can be used to identify underperforming research areas and highlight priorities for undertaking research to understand why *ex post* impact is not meeting expectations in those areas.

Key references
Raitzer D. and Kelley T.G. 2008. Benefit–cost meta-analysis of investment in the international agricultural research centers of the CGIAR • A condensed version – focusing on the 'Big 5' success stories of the Raitzer (2003) meta analysis – that uses impact accounting with success stories to systematically evaluate the returns to the total investments in the CGIAR.

Walker T. and Crissman C. 1996. Case Studies of the Economic Impact of CIP-Related Technologies. • One of the first applications of impact accounting with success stories at the level of a research institute.

Programmatic epIA

The CGIAR uses several mechanisms to review its programs. Almost all centers have conducted center-commissioned external reviews (CCERs) periodically on selected programs. Centers and their programs are also reviewed every five years in external program management reviews (EPMRs). Information from epIAs potentially feeds into both these types of reviews. The issue at hand is not whether or not programmatic reviews should take place, but whether or not programmatic epIAs are desirable.

In Chapter 3, the conduct of epIA is discussed from the perspective of research that generates a well-defined technology or policy resulting in a success story for documentation in a Stage I epIA. Such success stories accumulate over time. EpIA can also be carried out at the program level in retrospective analyses recording the technological change that a research program has contributed to. Note that programmatic epIA does not refer to the meta-analysis of a large research area across research institutes but pertains to epIAs conducted at the level of research areas (programs) within a research institute. In the CGIAR, programmatic epIAs have generally focused on genetic improvement programs (Evenson and Gollin, 2003)[5].

In international agricultural research, a programmatic epIA is a major undertaking and may take several years to complete. Resources have to be mobilized and many scientists (usually 50–100 from many partner institutions) participate in supplying information that almost always blends harder field estimates with softer expert opinion (Evenson and Gollin, 2003). Presumably, each succeeding programmatic evaluation should be easier to carry out than the preceding one. In the past, results from such evaluations of plant breeding programs have provided valuable information on accountability and have contributed to learning about how to make programs more effective. In the future, periodic programmatic evaluations should be able to furnish information on the returns to agricultural research over time across the mandated commodities of the CGIAR. Indeed, documenting the returns to agricultural research over time should be the overriding concern in such reviews, particularly now that a baseline has been established in Evenson and Gollin (2003). Every ten years seems like a reasonable interval for conducting an epIA of a genetic improvement program.

EpIA practitioners who are assigned the arduous task of orchestrating these reviews always face the temptation of stopping

5 A recent and interesting example of a programmatic impact assessment saw ACIAR undertaking the epIA of a cluster of its animal health projects (ACIAR, 2006). Although benefits and costs were not updated and aggregated across the program, earlier *ex ante* analysis and expectations on each project were confronted by their *ex post* reality.

short of implementing an assessment equivalent to a Stage I epIA. Adoption outcomes are easier to report than estimated benefits spread across multiple countries. Benefit estimation is taxing in a programmatic epIA, but the rewards gained from making the extra effort to generate order of magnitude estimates can help to locate the performance of CGIAR's genetic improvement program in a spatial, temporal, and commodity context.

Another issue in programmatic epIA concerns its desirability and feasibility in research areas that are not related to plant breeding *per se*. Meta-analyses, such as Alston et al. (2000) and Evenson (2001), of epIAs for agricultural research show that the genetic improvement of major field crops accounts for 85% of the studies, with the CGIAR being no exception (Raitzer, 2003). Aside from the biological control of exotic pests and IPM, a programmatic epIA would not be an effective tool to elicit information on impact because sufficient Stage I epIAs have not accumulated over time in the non-genetic improvement areas. For these research areas, Stage I epIA is still an important priority. Most plant breeding programs in the CGIAR have demonstrated conclusively that they have more than paid for the resources invested in them. Periodic programmatic epIA would appear to be the priority for these programs. Stage I epIA should be earmarked for analyzing the effects of novel cultivars, such as those first produced by marker-assisted selection, transgenic transformation, or participatory plant breeding; or cultivars that are characterized by qualitatively different traits from earlier varietal change attributed to the program. Some plant breeding programs have yet to establish their *bona fides* in filling an impact profile. For these programs, Stage I epIA continues to be a priority.

Good practice 2.14. Research areas with a good record from economic rate of return assessments should engage in programmatic epIAs. Multi-dimensional impact assessments are also a priority for those areas with a well-documented track record for generating practical impact. For research areas where sufficient success has yet to be recorded, economic rate of return assessments remain the priority.

Good practice 2.15. Given the effort that programmatic epIA entails, ten years seems to be an adequate interval between programmatic epIAs for areas of research such as plant breeding which have substantially demonstrated impacts in the past.

Good practice 2.16. Although adoption data are the most important information conditioning impact in programmatic epIAs, impact practitioners are encouraged to engage in assessing multi-country or multi-regional economic rates of return and the estimation of benefits.

Key references

ACIAR (Australian Centre for International Agricultural Research). 2006. Review of ACIAR-funded Animal Health Research •An insightful confrontation of *ex ante* expectations with *ex post* reality in a cluster of research projects.

Evenson R.E. and Gollin D. (Eds) 2003. Crop Variety Improvement and its Effect on Productivity: The Impact of International Agricultural Research. • Comprehensively documents varietal change attributed to NARS and 13 plant breeding programs in the CGIAR.

3. Conducting EpIA

This chapter identifies the principles of good practice for conducting epIAs. These principles refer to expectations on the availability of information that conditions transparency and analytical rigor. The presentation is tailored to Stage I-type disaggregated economic rate of return assessments. However, where appropriate, expectations about the availability of information pertinent to multi-dimensional impact assessments (especially of the Stage II type) are also discussed.

The discussion is organized roughly in the chronological order of activities that take place in a Stage I epIA. The selection of the research-derived technology or policy for epIA (in a well-defined institutional setting) is the subject of the first section.

The second section considers how adoption is a necessary condition for the impact of technology-oriented epIAs. The same section then looks at policy-oriented epIAs and discusses how documenting uptake and influence are key ingredients for making a persuasive case for attributing behavioral change to research. This section examines in depth the adoption and diffusion of technologies and the uptake and influence of policies. Reliable documentation on these important topics is a prerequisite for persuasive epIA.

The third section describes the counterfactual scenario, which is the center of attention in the general literature on impact assessment, along with its good practice application in epIAs of agricultural research.

The estimation of benefits and costs and the presentation of results are discussed in the fourth and fifth sections respectively. The fifth section ends with extensive coverage of how epIAs can inform about a program's effect on alleviating poverty, which is the apex impact in Figure 2.1 and one of the MDGs that attracts the most interest from donors (Raitzer and Winkel, 2005).

Selecting the research-related technology or policy for epIA

The selection of research-related technologies for epIA is constrained by funding and by the number of perceived success stories at any point in time. Some research-related technologies that build on previous work may have special project (restricted) funds for epIA. But, in general, funding for epIA is scarce. A typical scenario is that economic rate of return assessments have to be supported by unrestricted core funds, whilst special funding usually needs to be solicited for more costly multi-dimensional impact assessments. A scarcity of funds, particularly for operational expenses, argues for greater scrutiny in the selection of research-derived technologies and policies for epIA.

The previous chapter endorsed the success story approach for general accountability purposes. To qualify as a success story, a research-related technology should be characterized by a level of adoption that has the potential to generate sufficient benefits to at least cover the costs of the research on technology generation and transfer. Therefore, evidence of adoption is the primary consideration in determining the suitability of candidate technologies and policies for impact assessment. At the selection stage, adoption evidence tends to be largely anecdotal and may be contested. Candidate technologies with anecdotal evidence should be screened for their desirability and feasibility for epIA.

The first step in an illustrative protocol for screening candidate technologies on the grounds of their desirability for epIA is described in Figure 3.1. This figure sets out three minimal thresholds that should be passed by the prospective epIA technology before advancing to the next step, which consists of a desirability scorecard.

Component novelty

That a technology should be novel to prospective adopters appears so self-evident as to seem trivial. Although the criterion of