Available online at www.sciencedirect.com
SCIENCE @ DIRECT®

ELSEVIER

Agricultural Systems 78 (2003) 166–184

AGRICULTURAL SYSTEMS

www.elsevier.com/locate/agsy

Why impact analysis should not be used for research evaluation and what the alternatives are

Javier Ekboir*

CIMMYT Economics program, Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico

Abstract

Many impact studies relate changes in impact indicators to research investments. This is valid only if an implicit assumption is true: that the link between indicators and investments dominates all other relationships that influence the impact indicators. However, this is only true for minor improvements along stable technological paths. In most cases, other factors, such as policies and markets, influence adoption and, consequently, impact. The problem is compounded because impacts often appear after many years and usually cannot be measured. Since many factors influence adoption, research impacts should be analyzed as part of a complex adaptive system that depends on external forces (e.g., markets), the direct and indirect interactions among agents (e.g., researchers, input suppliers and farmers), and the technology's nature and evolution. The complexity framework has broad consequences for agricultural and research policies. Since impacts result from the actions of the whole network, they cannot generally be attributed to individual agents. In evaluating networks, the relevant parameters to study are the rules for generating, collecting and sharing information, financing procedures, intellectual property-rights regulations and availability of human and financial resources. For individual agents the relevant indicators are their patterns of participation in particular networks, benefits and costs of participation, evaluation criteria, financial arrangements and institutional cultures.

Keywords: Complexity; Research policy; Random processes

1. Introduction

Several studies have estimated the impact of research investments through changes in social welfare or agricultural production; in some cases, impacts have also been

---

* Tel.: +52-5804-2004x2202; fax: +52-5804-7558/59.
E-mail address: j.ekboir@cgiar.org (J. Ekboir).

doi:10.1016/S0308-521X(03)00125-2

used to estimate research efficiency (Alston et al., 1995; Brennan, 1999; Davis, 1994; Jain and Byerlee, 1999; Maredia and Byerlee, 1999). As will be shown in this paper, this procedure is flawed because:

(a) impacts are the consequence of research outputs interacting with many variables that influence adoption (e.g. the technology's intrinsic characteristics, adopters' features, the effectiveness of extension services, markets, policies, regulations and globalization) within a system characterized by multiple interactions among several agents and institutions; in other words, impacts (or, more commonly, lack of) may result from a number of causes entirely unrelated to the quality of the research being evaluated;
(b) in such a complex system, impacts result from the interaction of several causes preventing the attribution to individual variables;
(c) impact assessment depends crucially on the measurement of research inputs and outputs but no adequate measures of these have been developed yet; in some cases, impact assessment also depends on the postulated relationship between research inputs and impacts; and
(d) the definition of what constitutes the "technology being adopted" is usually arbitrary because research outputs often are further developed by users during adoption; this problem is less important for embedded technologies that require only minor changes in prevalent practices.

Since impacts occur within complex adaptive systems (CAS)[1] that involve many agents interacting both randomly and purposively, the use of impact assessment to define scientific policies without analyzing the processes that generated the estimated results is equivalent to confusing the outcome of a random experiment with the underlying random process. Even though the outcome provides some information about the process that generated it, the latter can only be understood with a complete characterization of the variables that determine its time path and the underlying probability distribution of the stochastic processes. Similarly, impact assessment provides little information if the research and diffusion process that generated it is not understood.

Seeing research as part of a CAS highlights the fact that the use of impacts as a guide for research policies would be valid only if the following five assumptions held: (1) there is a direct causal link going from research to impact indicators; (2) this link dominates all other relationships and variables affecting adoption; (3) chance has no influence on the relationship between investments in research and impacts; (4) the first two assumptions are valid over the whole period that goes from the start of research to measurement of its impacts; and (5) inputs and impacts can

---

[1] A complex system is defined as one that cannot be understood by the analysis of its isolated components (Gallagher and Appenzeller, 1999). The system is adaptive if its evolution is influenced by its history (Kauffman, 1995). Two important features of CAS are that chance usually plays a role in the system's evolution and that there are many direct and indirect interactions among the system's components.

be measured; in the case of ex ante assessments, impacts can be predicted to an acceptable degree of accuracy.

As will be shown in this paper, the five assumptions are seldom valid and when they are, it is only for minor changes along stable technological paradigms (such as successive replacements of modern varieties). The first three assumptions fail because technology adoption and impact are components of a CAS (see Sections 2 and 3). The fourth assumption usually does not hold because the nature of these interactions changes as the technologies and markets mature (Rycroft and Kash, 1999). The fifth assumption will be discussed thoroughly in Section 5.

The fact that a technology had no impact could be completely unrelated to the quality of the research that originated it. Conversely, a technology can have an impact even though it was not developed by research institutions. The problem is essentially how to define causality in CAS where many agents and processes interact through multiple channels and feedback loops. When the first four assumptions hold, the system is simple and the link between investment and impacts can be described by a stable, simple mathematical relationship. In a CAS, though, several variables and chance interact to produce the observed results, making it impossible to assign causality to just one variable as the process depends on the whole set of variables and their interactions.

The approach presented in this paper is a significant challenge to the dominant discourse on impact assessment, which assumes that agents can manage well defined processes. The idea of causality in CAS, on the other hand, stresses how limited that ability is since each agent interacts with many others, together producing the emergent dynamics of that interaction. Each agent will make choices, trying to influence outcomes and the dynamics in which those outcomes emerge. What finally emerges will be a result of the conflicting constraints that the environment and the various agents place on each other, and not the result of the choices of individual agents (Stacey et al., 2000).

Section 2 discusses CAS's features relevant for the analysis of research systems. Section 3 defines the National Innovation System: a system that includes not only researchers but all agents that influence the adoption of innovations and their patterns of interaction. The relationship between the Innovation and Research Systems is discussed in Section 4. Section 5 discusses the issues involved in the measurement of impacts and research inputs and outputs. Using the concepts presented in the previous sections, Section 6 analyzes the reasons for the perceived limited impact of research systems in developing countries. Sections 7 discusses alternative indicators for the assessment of research activities. Section 8 contains the concluding remarks.

2. Science and technology as a complex adaptive system

The traditional analysis of technical change considers technology as a black box that shifts the production function, but the evolution of the box itself is an issue beyond economic analysis (Nelson and Rosenberg, 1993). Technical change has usually been represented as a mechanism, i.e., the whole system could be understood by the analysis of its parts and could be described by well defined deterministic equations always moving towards equilibrium. When uncertainty was introduced, it could be reduced with better information. All causes could be separated and analyzed individually, using the assumption of caeteribus paribus. This mechanistic vision has lead to an emphasis on predicting the future, measuring activities and exercising strict control (Stacey et al., 2000).

An alternative framework for the study of technical change is based on evolutionary economics, institutional economics, stochastic processes and theories of complexity (Arthur et al., 1997; Dixit and Pindyck, 1994; Kauffman, 1995; North, 1991; Rycroft and Kash, 1999; Stacey et al., 2000). Key features of this framework are:

- Agents have limited information and understanding of the environment in which they live. Therefore, agents are not assumed to maximize profits but to use bounded rationality.
- A minority of agents is always looking for new technological and institutional opportunities, some of which will eventually be adopted by the majority of agents.
- These processes self-organize by the interactions of many agents who follow their individual plans. Even though some agents have more clout than others, no single agent or group of them has the power to determine uniquely the development path.
- CAS evolve by the interaction of trends and random events, subject to the initial conditions. The distribution of the random events is highly skewed; in other words, these processes evolve through a succession of many small variations interrupted by rare catastrophic mutations. The latter can be triggered by small changes in any variable and spread through the system (Kauffman, 1995). Even though it is possible to model the probability distribution, it is impossible to predict whether the next change will be small or catastrophic, i.e., CASs are path-dependent and unpredictable in the long term. Limited predictability (especially of major trends) is possible, but random events may derail these predictions. Additional information can reduce but not eliminate the uncertainty which increases with the time horizon considered (Dixit and Pindyck, 1994).
- From the previous paragraph it follows that CAS are not completely random. Even though particular outcomes cannot be predicted, it is possible to study the nature of their trajectories. Of special interest for policy analysis, it is possible to analyze the emergence of aggregate behaviour resulting from the actions of agents interacting in a particular environment. This analysis can indicate which policies and instruments have the highest probability of influencing the system in a desired direction.

One key feature of this approach is that the profitability of any technology depends not only on its intrinsic characteristics and its own market but on those of complementary and substitutive technologies (Kauffman, 1995). Research, new technologies and economic variables change a technology's profitability directly and indirectly. The direct effect is the change in the technology's own profitability, for example, as a result of improvements to it. The indirect effect is the change in the technology's profitability induced by variations in other technologies. For example, typewriters became obsolete after the massive adoption of the personal computer. The quality of research on typewriters was absolutely unrelated to their technical obsolescence. Also, the diffusion of the personal computer opened the market for a whole range of peripherals.

Economic agents use their unique abilities, learning procedures and incomplete information to search for new technological opportunities in this evolving technology space (Kauffman, 1995; Rycroft and Kash, 1999). In other words, technical change results from the exploration of unknown areas of the technology space. But adoption of a new technology changes the technology space locally and non-locally through the direct and indirect effects mentioned above. Because the simultaneous exploration of the technology space by a multitude of agents deforms the same space they are exploring in ways that cannot be predicted, all learning is local and short lived. Past events can be described in detail, but that will not help to predict how the technology space will evolve (Kauffman, 1995; Dixit and Pindyck, 1994).

Even though many developments are unpredictable, sometimes it is possible to identify emerging technological trends. However, these trends can be understood fully only after they have emerged; they cannot be predicted entirely in advance. Early identification depends on the information flows within the system and the individual and collective ability to process it. Making strategic decisions is equivalent to positioning an institution along a technological trajectory and to defining procedures to react to unexpected changes in the technology space. The better a trajectory is understood, the greater the benefits that can be achieved (Rycroft and Kash, 1999).

3. National innovation systems

In contrast to the simple mechanistic approach to technical change, a new approach, based on the CAS framework, has been developed in the last three decades. The key concept in this framework is innovation, defined as anything new introduced into an economic or social process (OECD, 1999; Archibugi et al., 1999).

Essentially, innovation is the ability to manage knowledge creatively in response to market-articulated demands or other social needs. Knowledge flows and their transformation into innovations depend on the intrinsic characteristics of knowledge, the legal system, formal and informal regulations, and by the process's history. The *national innovation system (NIS)* is composed by all agents involved in the innovation process, their actions, interactions and the formal and informal rules that regulate the system (OECD, 1999). Six characteristics of the NIS relevant to this paper are:

1. Knowledge creation is essential to the NIS's dynamics. Learning occurs at all levels: individuals, institutions and society as a whole (Archibugi et al., 1999).

2. An innovation does not have to be new for the world or even the country where it is adopted, but only for the agent that adopts it (OECD, 1999). While the traditional approach emphasizes novelty for the world, the NIS approach emphasizes novelty for the adopting agent.
3. The dynamics of the NIS do not depend on the agents at the forefront of research and technology development, but on the innovative capabilities of the majority of agents. In other words, it is more important to have many agents searching for and adapting existing technologies than to have a few sophisticated research institutes in a static society.
4. Each agent's innovative abilities depend crucially on the information flows within the NIS and on its ability to process the information (Salter and Martin, 2001). Stronger flows enable earlier identification of technological and market opportunities.
5. In general, adoption of innovations, including embodied technologies, require organizational changes within the firm (Rycroft and Kash, 1999). In other words, innovations cause both shifts along the axes and changes in the production function itself.
6. Technologies are increasingly generated and disseminated by networks of agents who interact directly and indirectly (through markets and other channels). The nature of these interactions depends to a certain extent on the complexity of the technology and its maturity. Simple mass production technologies are basically produced by isolated teams of researchers working in one institution (usually a laboratory within a university or firm) while complex technologies are produced by networks that co-evolve with the technologies they generate. The networks involve not only researchers from different institutions but users of technology, input suppliers, government agencies, NGOs and/or financial institutions. Indirect interactions (especially market interactions) are more important for simple or more mature technologies because each agent understands the needs and roles of other agents in the network. On the other hand, newer or complex technologies require closer interactions due to the greater uncertainty about technical standards, users' needs and markets (Rycroft and Kash, 1999).

4. The innovation and research systems

The traditional model of science describes the generation and adoption of new technologies as a continuum that starts with basic science, continues with applied research and ends with technology transfer. This linear vision does not describe most historical experiences, where technology development preceded the scientific understanding of the processes behind them, as exemplified by the steam engine and thermodynamics (IDRC, 1997; Nelson and Rosenberg, 1993). In the NIS framework, technologies are developed by networks of agents that feature several feedback loops. More than a linear process, innovation processes resemble a spider web.

The traditional vision of science has been behind the use of impact assessment for the definition of research policies. In this framework, more inputs in the first stages of the process result in more technologies after a certain period of time, justifying the connection between investments in research and impacts. This vision has also influenced the trend observed in many public research institutions towards increased use of more formal priority setting mechanisms and greater control of researchers' activities (Ekboir, 2001).

The causal link between research inputs and impacts breaks in the framework of the NIS, because the latter is larger and more complex than the research system. Several reasons justify the analysis of the broader system:

1. Commercial firms are the most important source of innovations: their success depends on their ability to take advantage of economic and technological opportunities (OECD, 1999).
2. Innovative activities are interactive processes that involve market and non-market institutions.
3. Most research outputs are modified when they are incorporated into production processes (Ekboir, 2001; Morris et al., 1999; Smale at al., 1999). These modifications generally result from non-formal research.
4. A NIS may be efficient even if the research system is weak and inefficient, as exemplified by the Italian experience after WWII (Malerba, 1993). Conversely, a research system may be strong while the innovation system is weak, as in the Soviet Union of the 1970s.
5. Technologies are becoming more complex not only because they require more components and a stronger scientific base but also because globalization and technical change are forcing agents to compete in faster changing markets. Products that were produced with simple technologies and sold in stable local markets, are increasingly integrated into the international economy, as exemplified by basic grains in Central America. Even if agricultural practices do not change, farmers are forced to deal with new marketing challenges that demand new technologies and skills. Due to the greater complexity, technologies are increasingly developed and diffused by networks (Rycroft and Kash, 1999).

5. Measuring efficiency and impact

As explained above, the mechanistic vision of research has framed the discussion of scientific policies and shaped research evaluation procedures, including the use of impact analysis as an indicator of research efficiency. In addition to the indicated attribution problems, the use of impact assessment for research evaluation is hampered by measurement problems.

A distinction has been made between impact or efficacy—the capacity to produce effects—and efficiency-the ratio of output to input (Anderson and Hardaker, 1992;

Alston et al., 1995). Many impact studies directly linked efficacy with efficiency by estimating the economic impact of research investments. As was seen above, this link is valid only under restrictive assumptions.

Estimation of research efficiency requires (a) identification and measurement of impacts, (b) measurement of research inputs and outputs, (c) modelling the process by which the research inputs effect impact indicators, and (d) attribution of impacts to the identified causes. We will now review these issues. The main conclusion is that impacts cannot be identified for research institutions in isolation but only for the NIS, while efficiency cannot be estimated at all.

5.1. Research outputs cannot be measured

It has not been possible to define measures of research outputs common to all categories of science, because each field has its own particular outputs and procedures for disseminating them (OECD, 1997; Okubo, 1997). For example, bibliometric analysis can be used in physics, chemistry and biomedicine because they are well represented in the SCI, but geosciences, biological field research, mathematics, engineering and technology are not. In some institutions (e.g., universities) the main research output is a journal article, while for others in the same field (e.g., laboratories in the private industry) the main output is a patent (Stoneman, 1995).

Often it is impossible to assess the timing and magnitude of the impacts. For example, the structure of DNA was first published in 1953; until the late 1980s there were no commercial applications for this knowledge. An ex post evaluation of investments in biotechnology in 1987 would have shown very little impact. Today, their total impact cannot be measured because new applications are constantly being discovered. Precise ex ante measurement is even more difficult, except for minor developments along established research lines. For example, the first paper on lasers was rejected for publication. But even in incremental research, unforeseen products with high payoffs can be discovered; for example, Viagra was found while searching for a heart medication. Despite the difficulty of precise ex ante measurement, biotechnology is relentlessly pursued because of its enormous current and future economic implications.

Until now it has been impossible to define good measures of innovative efforts (which include tangible outputs, design improvements and organizational changes) or of the inputs used in them. For example, many informal trials are financed and conducted by farmers in their farms and by agrochemical companies as part of their promotion strategies. These resources are not included in the official data of research investments. The importance of these resources is increasing with the greater use of participatory research methods.

High-quality research activities may or may not produce an output, but they always produce knowledge. Even though the output may have no impact, the knowledge produced during the research process is valuable because it could be used by other researchers to focus their activities, increasing the probability of success of other research projects. In the economics literature, these potential benefits are known as an "option value". Despite acknowledgement of the option value of knowledge, it has never been formally included in impact assessments.

5.2. Measures of research inputs are also ill-defined

In addition to problems similar to the ones identified above, measurement of the quantity and quality of research inputs presents two additional problems. First, it is very difficult to measure the effort made by researchers. Because of the uncertainty that characterizes research, especially the most novel or risky fields, the quality of a researcher cannot be inferred from a failure to obtain results (Geroski, 1995; Huffman and Just, 2000).

Second, the quality of a research input may depend on its interaction with other inputs. This problem especially affects researchers as it has been shown that their productivity depends heavily on the environment in which they work (Hare and Wyatt, 1988; Anderson and Hardaker, 1992). In other words, scientists tend to conform to the demands of their institutions. Then, more than inadequate capabilities, a low productivity may reflect management problems, low standards or insufficient resources to work with. For example, about 90% of the professionals of some Latin American research institutions have formal graduate training but they do not have support personnel to work with. Consequently, a PhD may end up doing the work of a laboratory assistant or doing no work at all.

5.3. Linking research inputs to outputs

Assuming that measurement issues could be solved, the estimation of research efficiency also requires establishing the nature of the relationship between inputs and outputs. This relationship can be defined as the ‘‘research production function’’ and, in general, it has not been derived from assumed maximizing behaviour. In particular, it is important to determine whether there are any economies or diseconomies of size or scope, feedback loops, spillovers and institutional factors including adequate and stable levels of funding and proper incentives for researchers.

The nature of the research production function for private companies was analyzed, among others, by Henderson and Cockburn (1996), Klette (1996), Schmookler (1972) and Zenger (1994). However, very few studies have been conducted for public institutes. Branson and Foster (1987) found that the cost function at USDA experimental stations is U-shaped, meaning that economies of size exist for smaller units while larger units face diseconomies of size.

Knowledge of the properties of the research production function is essential to measure efficiency. For example, several studies have observed that wheat germplasm flows mainly from larger breeding programs towards smaller ones, concluding that major economies of size exist in breeding (Brennan, 1999; Byerlee and Traxler, 1996; Jain and Byerlee, 1999; Traxler et al., 1995). The policy recommendation was that the efficiency of breeding programs could be increased by reallocating resources to the programs that generate spillovers.

The anecdotal evidence, however, can be misleading for the observed pattern could have several causes (i.e., economies of size, economies of scope, feedback loops within the production of knowledge, spillovers or institutional factors) and different policy recommendations would be derived from each one. For example, if economies of size are important, larger programs should be strengthened. However, if the research production function is characterized by diseconomies of scale and the low productivity of smaller programs is due to their lack of resources, the correct policy is to strengthen the latter. Finally, if low productivity is the result of inadequate incentives offered to researchers or unstable funding, the right policy would be a series of institutional reforms. In the last two cases, implementation of the right policies would create bi-directional flows, increasing the efficiency of the whole system.

An additional reason that hampers the estimation of research efficiency is that large variations in the productivity of research institutions within the same country have been identified, suggesting that idiosyncratic institutional and locational factors are responsible for these differences. A number of studies (Byerlee and Alex, 1998; Collins and Yeager, 1988; Hare and Wyatt, 1988; Holmstrong, 1989; Huffman and Just, 2000; IDRC, 1997; McClellan and Dorn, 1999; Rozelle et al., 1997; Zenger, 1994) show that:

1. There is a great disparity between the productivity of individual researchers.
2. The quality of researchers depends on the incentives structure, procedures for hiring and firing personnel, rotation of researchers, level and variability of salaries and operating budgets, and the history of each research group.
3. Evidence on the link between funding levels and the quality of scientific research is scant.
4. Both the quality and the quantity of research results are strongly influenced by researchers' perceptions of the quality of the institution's (and the team's) management.

Omission of these factors in the analysis of research efficiency would result in spurious regressions or biased impact estimates.

5.4. *Attributing impacts to different causes*

Because of the interactions between self-organization, trends and chance, complex adaptive systems evolve through the interplay of stability and change. The traditional analysis of technical change has focused on stability, separability of causes and control. In this framework, agents only interact through stable and predictable markets. Complexity theories, on the other side, emphasize multiple market and non-market interactions, change and unpredictability. Because of these interactions, most causes cannot be separated from each other. In other words, causality can only be attributed to the whole set of causes and not to its individual elements.

Research impacts result from the transformation of research outputs into innovations. Thus, the lack of impact of a research system may result from its own inefficiency, from failures in the transformation of research products into technologies, from problems in other areas of the NIS (such as lack of input markets or adverse policies) or from developments in related markets or technologies. On the other hand, a technology may have a major impact even though the national research system had little input in its development.

In addition to the problems explained above, ex ante estimation of impacts also requires advanced identification of the technology that will provide the best solution; however, as was seen above, it is not possible to know in advance which technology will be adopted. These points can be exemplified with the interaction between plant breeding and zero tillage in the Mercado Común del Sur (MERCOSUR). Usually, the impact of plant breeding has been measured by welfare gains or increases in grain production (Alston et al., 1995; Brennan, 1999; Davis, 1994; Jain and Byerlee, 1999; Maredia and Byerlee, 1999). This methodology would be valid only if seeds were the only change in the production function or if the shift of the production function induced by improved seeds could be measured separately from other changes.

Production of the five major grains in MERCOSUR (maize, sorghum, sunflower, soybeans and wheat) jumped from 23 million tons in 1961 to 152 million tons in 2001 (FAO, 2001). This increase resulted from higher yields and an expansion of the cultivated area, both enabled by improved germplasm and new crop management technologies, especially zero tillage. The introduction of soybeans in the late 1960s induced an intensification of agriculture that caused serious soil degradation in the traditional cropping regions. Zero tillage reduced production costs, reversed soil erosion and allowed an expansion of agriculture into previously marginal lands. Without zero tillage, grain production would have had to be abandoned in many places (Ekboir, 2001).

While plant breeding partially fits the linear vision of science, the development of zero tillage does not. The first sustained efforts to develop zero tillage were conducted by ICI-an agrochemical company- willing to create a market for a newly developed herbicide. In Brazil, ICI teamed with a few researchers from public institutions and from an international cooperation agency, farmers and equipment manufacturers. In Argentina, the technology was developed by individual public researchers and farmers interacting with equipment manufacturers. In Paraguay, zero tillage was developed mainly by farmers with connections to Brazil, supported by an international cooperation agency. In all countries, the development effort took about 15 years; by 1985 a technically efficient package was available. Adoption, however, remained low until the early 1990s because a crucial input (the herbicide glyphosate developed by Monsanto) was expensive. When the price of glyphosate fell from 40 US$/l to less than 10 US$/l, zero tillage became economically efficient and adoption exploded (Ekboir, 2001).

The impact of these events cannot be separated. Without zero tillage, the impact of improved germplasm would have been very small as zero tillage was necessary to stop soil erosion and improve water management. At the same time, the improved germplasm increased the profitability of zero tillage, fostering adoption. But adoption only exploded when a key input produced by a private firm became affordable. How should the impact be allocated to the components of the package: privately developed herbicides, publicly or privately developed seeds and the zero tillage package developed by a network of agents? Any division of the impact between these factors would only be a non-robust computational result, with little practical meaning. How could research be evaluated by the impact of the package? Without the reduction in the price of glyphosate the impact would have been small; with the price reduction, it was significant. But the price reduction was completely unrelated to plant breeding or development of zero tillage.

It was mentioned that ex ante impact assessment has to decide in advance which technologies (or families of) would be adopted. In the twenty years required to develop zero tillage in Brazil, three alternatives were developed. Two of them, based on terracing, were inferior to zero tillage and were not sustainable in the long term. However, policy makers used ex ante assessments to select the inferior alternatives as part of soil conservation programs. Only after these technologies proved to be unsustainable, was zero tillage recognized as the best option.

6. Why has agricultural research in developing countries not had a greater impact?

Active innovation processes have been associated with flexible and active networks (Ekboir, 2002; Rycroft and Kash, 1999). However, this has not been recognized in most developing countries where the organization of agricultural research and extension has been greatly influenced by the linear vision of science. Thus, these activities were the responsibility of specialized institutions; when they were in the same institution, it was under the responsibility of separate administrative units.

Researchers were supposed to create knowledge within the research institutions, with little interaction with extension agents, private firms or farmers. Seen as the creators of knowledge, researchers often established a hierarchical relationship with extension agents; this hierarchy also characterized the interactions between extension agents and farmers. In many cases, research institutions were created to support specific public agencies, with little incentives to interact with the private sector. The unidirectional information flows distorted the perception that researchers and policy makers had of farmers' needs, resulting in the development of technologies that were never adopted and top-down transfer programs. In addition, in most public institutions there were no quality standards for researchers or there were no procedures to enforce them.

This structure survived until the mid-1980s because most countries sought food self-sufficiency with little consideration for competitiveness or sustainability. In developing countries research was concentrated in a few export commodities, basic grains or livestock. These products had well established markets and production structures. As the scientific research on these products had developed over many decades, most agents knew how and where to look for information. The stable technology trend minimized the need for direct interactions. Support policies helped farmers to maintain non-competitive production structures. The lack of quality standards for research and extension allowed many professionals to produce low quality works.

After the collapse of the import substitution model, competitiveness and sustainability became major requirements of agricultural systems. At this point, many agents in the innovation system started to question the effectiveness of research institutions and their weak interactions with the private sector. The demand for impact assessment reflected the perceived lack of communication between the private and public sectors and the lack of quality standards. In response to the new environment, most developing countries' research institutions went through major reorganizations aimed at adapting them to a globalized environment and a more demanding private sector.

The main trends of research policies in the 1990s were reductions in public funding, diversification of its sources (with increasing reliance on competitive programs and sales of goods and services), increasing reliance on formal priority setting mechanisms (with particular attention to identification of technology demands) and increased requirements to show impacts. On the other hand, incentives and quality standards within the institutions changed little, except for the reduction in public funding, the requirement for researchers to generate their own operating funds and real wage reductions. The result was a greater awareness by managers of public research institutions of the need to interact with other agents within the NIS (but these interactions were seen mainly as a source of funds and not as true partnerships), greater uncertainty about funding for long term research programs and deterioration of the research infrastructure (Echeverría, 1998; Rozelle et al., 1997). A few countries (e.g., Argentina and Brazil) introduced incentives to foster interactions between public research institutions and private firms, but these efforts were hampered by lack of continuity and the increasing weakness of the public institutes.

7. Research policies in the NIS framework

The previous sections showed the shortcomings of the linear vision of science, including impact analysis, as a basis of research policies. This section discusses briefly the new perspectives for these policies provided by the CAS and innovation frameworks. The starting point is the recognition that (1) innovation processes are essentially uncertain, (2) since they self-organize and they cannot be controlled by any single agent, opportunities and challenges emerge continuously, and (3) impacts result from strong networks. These principles imply that the key issue is not managing the research process (in the traditional sense of managing stable industrial processes) but setting the appropriate conditions for the emergence of strong innovation networks and for the active participation of research institutions in them, in other words, creating an enabling environment for knowledge creation and sharing. The approach has concrete implications for the whole system, for institutions and individual researchers.

Since technology generation and adoption depend on the performance of the NIS, research policies should be part of broader innovation policies. The latter have been the subject of numerous works and will not be reviewed here (see OECD, 1999, for a thorough review of the issues and some experiences). We will discuss briefly only a few important issues.

Since research is a search in an ever changing knowledge space and the emergence of scientific and technological trends cannot be predicted, institutions should balance research along known lines with exploration of new ideas (Kauffman, 1995).

This search should be based on the principle that knowledge cannot be managed, only enabled through the introduction of appropriate sets of incentives and procedures to: (a) create and screen valuable ideas; (b) develop those ideas through disciplined project management; (c) provide leadership during the research process; and c) create a cultural environment conducive to innovation (Von Krogh et al., 2000). For this, research strategies and management should take an adaptive stance and be prepared to react to unanticipated events (Phelan, 1995). In recent years, a number of countries have developed "technology foresight" programs and methodologies to identify emerging trends in science and technology.[2]

Strategies should not be conceived as long-term guiding principles carved in stone but as permanent searches to identify emerging trends and driving forces. Traditional impact assessment is of limited utility because it focuses on specific outcomes at specific moments and not on a continuous monitoring of the processes that generated them. An essential tool for adaptive strategies is the establishment of adaptive and continuous monitoring and evaluation systems, such as those used in the last two decades to manage some complex environmental systems (Chess et al., 2000; Pulwarty and Melis, 2001).

Two essential issues in the design of a monitoring and evaluation system are what to monitor and how to set the evaluation standards. As was explained in Section 5 research outputs or impacts are poor indicators of the quality and intensity of research efforts or of the pertinence of research activities. In spite of these difficulties, persistent lack of impact of a research institution is an indicator that the institution is either not producing good quality research or is not interacting with its environment in a productive manner.

Although research inputs are also difficult to measure, the quality of research programs, their relevance and integration in the innovation system can be assessed through peer reviews and consultations with stakeholders and key partners in the innovation process. The composition of the panel is of extreme importance for the usefulness of the review process. Since the reviewers should not have a personal stake (either positive or negative) in the research being evaluated and should be acquainted with the latest developments in their areas of research, most developing countries should rely heavily on foreign professionals to integrate the review panel. These reviews are expensive and cannot be conducted on a yearly basis. Quantitative indicators of research activities and of networking (e.g., number of inter-institutional projects or number of papers co-authored with colleagues from other disciplines or institutions) can provide complementary information.

The definition of the indicators to be monitored should respond to the mandates of the research institutions. For example, since the 1990s several Latin American agricultural research institutions (e.g., in Brazil and Mexico) have based promotions on the number of papers published in international peer-reviewed journals. This has induced researchers to work in the research stations and to minimize interactions with other agents in the innovation system or with professionals from other

---

[2] The literature on technology foresight has exploded in recent years and its review is beyond the scope of this paper. A good introduction can be found in OECD (1996).

disciplines. Even though the number of publications has increased, the contribution of these institutions to the competitiveness of agricultural production has remained weak.

Emerging technological trends cannot be identified until they reach a minimum strength, thus, institutions can only establish research programs in fields already known. Also, efficiency indicates that institutions should only consolidate teams in areas that have already shown a strong potential to yield meaningful outcomes. On the other hand, since individual researchers explore less known areas, they can identify promising lines of research earlier; in fact, new trends can emerge only if researchers have the freedom to explore those less known areas.

Given the uncertainty about future trends and the best methods to approach specific research problems, it is important that networks and institutions maintain and support a plurality of ideas, including a certain amount of duplication of research efforts. Exploration of new research areas requires that personnel, financial and organizational slack is provided. Institutional performance is enhanced when uncertainty and instability are seen as the expected condition, and failures as essential to learning and rapid adaptation. Seeking perfect efficiency is the enemy of the slack needed to access and create the knowledge that will facilitate trend changes. In other words, researchers should have some freedom to conduct curiosity-motivated research.[3] But since curiosity- motivated research is more risky and less known, it cannot be evaluated with traditional management routines based on accomplishment of previously set objectives. (Huffman and Just, 2000; Stacey et al., 2000).

Strong interactions with private firms enable researchers and research managers to identify opportunities where they can contribute their technical expertise. In the 1990s, the emphasis has been on establishing joint financing mechanisms; these, however, do not contribute to the creation of strong work relationships between public researchers and private firm professionals nor to increase the information flows. More effective instruments to foster interactions are participatory research methods, setting experiments under production conditions outside the experimental stations, and joint R&D activities with commercial firms, in particular, public researchers spending time in the private firm's facilities and private professionals working in public research facilities.

Identification of research demands became a key component of research management in the last decade. In general, the methods used involved limited interactions with technology users. Also, after the demands were identified, there were no incentives to involve other agents in the actual research. But successful network adaptation requires more than just responding to technological demands. Successful networks must develop and maintain the ability to adapt to highly competitive environments in ways that also influence that environment. In other words, supply

---

[3] A discussion of the right incentives for researchers to balance the allocation of efforts between established programs and curiosity-motivated research is beyond the scope of this paper. It should be mentioned, though, that it is a key issue for the effective management of research institutions. For example, two problems that have to be solved are: should all researchers have the same freedom of research or only the most creative should be allowed to follow their interests? How to decide that a curiosity-motivated research effort should be upgraded to a full research program or should be completely abandoned?

of technology also has a major role to play in maintaining competitiveness. In general, supply factors will be more important in the early stages of technology development while demand will be more important in more established technological trajectories (Rycroft and Kash, 1999).

8. Final remarks

During the 1980s and 1990s substantial efforts were invested in the estimation of research impacts and the efficiency of research investments. These efforts reflected three prevalent perceptions among policy makers and economists: (1) socioeconomic processes functioned like mechanisms which could be understood by the study of its components and controlled with carefully designed policies; (2) following the linear vision of science, there was a direct link between the magnitude of research investments and research impacts, and (3) research institutions were self-centred and did not respond to the needs of other stakeholders.

As was shown, these efforts were the wrong answer to real problems. Since impacts result from the use of research outputs in productive or social processes, lack of impacts may result from a number of factors not related to the quality of the research being evaluated. Conversely, positive impacts result from effective interactions among a number of agents, which may not include researchers. Even though it is not possible to know in advance which particular research discoveries will yield important impacts, it is known that the strength of innovation networks and the quality of research institutions have a substantial influence on the probability of success. In other words, valuable information for the management of research systems cannot be obtained from one-time measures of outcomes but from continuous monitoring of the processes that produce the outcomes. Incentives offered to researchers and research administrators should be designed with these two factors in mind.

In spite of the difficulties to forecast particular technological outcomes, it is possible to identify emerging technological trends. These, however, cannot be identified until they have reached a minimum development. For this reason, institutions can only establish research programs in relatively known fields. On the other hand, because researchers explore less known areas, they can identify promising lines of research earlier; in fact, new trends can emerge only if researchers explore those less known areas.

Effective research management requires balancing the concentration of resources in established programs with granting researchers limited freedom to conduct curiosity-motivated research. But since curiosity- motivated research is more risky and less known, it cannot be evaluated with traditional management routines based on accomplishment of previously set objectives.

Finally, meaningful quantitative estimates of research efficiency are impossible to obtain due to problems related to (1) the definition and measurement of research inputs and outputs, (2) knowledge of the relationship between research inputs and outputs, (3) separating causes when several of them interact to produce an effect, and (4) ex ante estimation of impacts in a random setting. Because of these problems, estimates of efficiency are non-robust mathematical constructions that offer little guidance for innovation and research policies. Instead of impact indicators for the design of research policies, policy makers should demand the establishment of strong continuous monitoring and evaluation systems that track the quality of research programs and their interaction with other agents within the NIS.

Acknowledgements

An earlier version of this paper was presented at the international conference "Why Has Impacts Assessment Research Not Made More of a Difference?", held 4–7 February 2002 in San José, Costa Rica. The conference was co-organized by the Standing Panel on Impact Assessment (SPIA) of the Consultative Group for International Agricultural Research (CGIAR) and by the Economics Program of the International Maize and Wheat Improvement Centre (CIMMYT).

References

Alston, J., Norton, G., Pardey, P., 1995. Science under Scarcity: Principles and Practice for Agricultural Research Evaluation and Priority Setting. Cornell University Press.

Anderson, J., Hardaker, J.B., 1992. Efficacy and efficiency in agricultural research: a systems view. Agricultural Systems 40, 105–123.

Archibugi, D., Howells, J., Michie, J., 1999. Innovation systems and policy in a global economy. In: Archibugi, D., Howells, J., Michie, J. (Eds.), Innovation Policy in a Global Economy. Cambridge University Press, UK.

Arthur, W.B, Durlauf, S.N., Lane, D.A., 1997. Introduction. In: Arthur, W.B, Durlauf, S.N., Lane, D.A. (Eds.), The Economy as an Evolving Complex System II. Santa Fe Institute in the Sciences of Complexity, Addison Weasley Longman, Inc, MA USA.

Branson, I., Foster, P., 1987. Economies of size of USDA agricultural research stations. IEEE Transactions on Engineering Management 34 (3), 156–160.

Brennan, J.P., 1999. Efficiency in wheat improvement research: a case study of Australia. In: Maredia, M.K., Byerlee, D. (Eds.), The Global Wheat Improvement System: Prospects for Enhancing Efficiency in the Presence of Spillovers. CIMMYT Research Report No. 5, Mexico, D.F.

Byerlee, D., Alex, G., 1998. Strengthening National Agricultural Research Systems. Policy Issues and Good Practice, World Bank, Washington, DC.

Byerlee, D., Traxler, G., 1996. The role of technology spillovers and economies of size in the efficient design of agricultural research systems. In: Proceedings of the Conference Global Agricultural Science Policy for the Twenty-First Century, Melbourne, Australia, pp. 392–429.

Chess, C., Hance, B., Gibson, G., 2000. Adaptive participation in Watershed Management. Journal of Soil and Water Conservation 55 (3), 248–270.

Collins, R.A., Yeager, J.L., 1988. Staff evaluation and incentive practices utilized by behavioral science research organizations: a pilot study. Journal of the Society of Research Administrators 119–134.

Davis, J.S., 1994. Spillover effects of agricultural research: importance for research policy and incorporation in research evaluation models. In: Anderson, J.R. (Ed.), Agricultural Technology: Policy Issues for the International Community. CAB International, Wallingford, UK.

Dixit, A.K., Pindyck, R.S., 1994. Investment under Uncertainty. Princeton University Press.

Echeverría, R.G., 1998. Agricultural policy issues in Latin America: an overview. World Development 26 (6), 1103–1111.

Ekboir, J.M., 2002. Developing no-till packages for small-scale farmers. In: Ekboir, J. (Ed.), World Wheat Overview and Outlook. CIMMYT, Mexico, DF.

Ekboir, J.M., 2001. Sistemas de Innovación y Política Tecnológica: Siembra Directa en el MERCOSUR. In: Díaz Roselló, R. (Ed.), Siembra Directa en el Cono Sur. Uruguay, PROCISUR.

FAO, 2001. FAOSTAT, www.fao.org.

Gallagher, R., Appenzeller, T., 1999. Beyond reductionism. Science 284 (April 2), 79.

Geroski, P., 1995. Markets for technology: knowledge, innovation and appropriability. In: Stoneman, P. (Ed.), Handbook of the Economics of Innovation and Technological Change. Blackwell Publishers Inc, UK.

Hare, P., Wyatt, G., 1988. Modeling the determination of research output in British universities. Research Policy 17, 315–328.

Henderson, R., Cockburn, I., 1996. Scale, scope, and spillovers: the determinants of research productivity in drug discovery. RAND Journal of Economics 27 (1), 32–59.

Holmstrong, B., 1989. Agency cost and innovation. Journal of Economic Behavior and Organization 12, 305–327.

Huffman, W.E., Just, R., 2000. Setting efficient incentives for agricultural research: lessons from principal-agent theory. American Journal of Agricultural Economics 82 (4), 828–841.

IDRC, 1997 A Decade of Reform, www.idrc.ca.

Jain, K.B.L., Byerlee, D., 1999, Investment efficiency at the national level: wheat improvement research in India. In: Maredia, M.K. and Byerlee, D. (Eds.), The Global Wheat Improvement System: Prospects for Enhancing Efficiency in the Presence of Spillovers, CIMMYT Research Report No. 5, Mexico, D.F.

Kauffman, S., 1995. At Home in the Universe: The Search for the Laws of Self-Organization and Complexity. Oxford University Press, NY.

Klette, J., 1996. R&D, Scope economies, and plant performance. RAND Journal of Economics 27 (3), 502–522.

Malerba, F., 1993. The national system of innovation: Italy. In: Nelson, R. (Ed.), National Innovation Systems, A Comparative Analysis. Oxford University Press, NY.

Maredia, M.K., Byerlee, D., 1999, Toward efficient allocation of research resources in the presence of spillovers: lessons from wheat improvement research. In: Maredia, M.K. and Byerlee, D. (Eds.), The Global Wheat Improvement System: Prospects for Enhancing Efficiency in the Presence of Spillovers. CIMMYT Research Report No. 5, Mexico, D.F., Mexico.

McClellan III, J.E., Dorn, H., 1999. Science and Technology in World History. The John Hopkins University Press, Baltimore MD.

Morris, M.L., Risopoulos, J., Bak, D.L. 1999, Genetic Change in Farmer-Recycled Maize Seed: A Review of the Evidence, CIMMYT Economics Working Paper No. 99-07, CIMMYT, Mexico, D.F., Mexico.

Nelson, R., Rosenberg, N., 1993. Technical Innovation and National Systems. In: Nelson, R. (Ed.), National Innovation Systems, A Comparative Analysis. Oxford University Press, NY.

North, D.C., 1991. Institutions, Institutional Change and Economic Performance. Cambridge University Press, New York, NY.

OECD. 1999, Managing National Innovation Systems, France.

OECD. 1997, The Evaluation of Scientific Research: Selected Experiences, OCDE/GD(97)194, France.

OECD. 1996, STI Review, Special Issue on Government Technology Foresight. No. 17, March 12, 1996, France.

Okubo, Y., 1997, Bibliometric Indicators and Analysis of Research Systems: Methods and Examples, STI working paper 1997/1, OCDE/GD(97)41, OCDE, France.

Phelan, E.E. 1995. From chaos to complexity in strategic planning. Presented at the 55th Annual Meeting of the Academy of Management, 6–9 August, Vancouver, Canada.

Pulwarty, R.S., Melis, T.S., 2001. Climate extremes and adaptive management on the Colorado River: lessons from the 1997–1998 ENSO Event. Journal of Environmental Management 63 (3), 307–325.

Rozelle, S., Pray, C., Huang, J., 1997. Agricultural research policy in China: testing the limits of comercialization-led reform. Comparative Economic Studies 39 (2), 37–71.

Rycroft, R.W., Kash, D.E. 1999, The Complexity Challenge: Technological Innovation for the 21st Century. Science, Technology and the International Political Economy Series, A Cassell Imprint, NY.

Salter, A.J., Martin, B.R., 2001. The economic benefits of publicly funded basic research: a critical review. Research Policy 30, 509–532.

Smale, M., Aguirre, A., Bellón, M.R., Mendoza, J., Rosas, I.M. 1999, Farmer Management of Maize Diversity in the Central Valleys of Oaxaca, 1998 Baseline Socioeconomic Survey, CIMMYT Economics Working Paper No. 99-09, CIMMYT INIFAP, Mexico, D.F., Mexico.

Schmookler, J., 1972. The size of firm and the growth of knowledge. In: Schmookler, J. (Ed.), Patents, Invention, and Economic Change. Harvard University Press, Cambridge, MA.

Stacey, R.D., Griffin, D., Shaw, P., 2000. Complexity and Management. Fad or Radical Challenge to Systems Thinking?. Routledge, NY.

Stoneman, P., 1995. Introduction. In: Stoneman, P. (Ed.), Handbook of the Economics of Innovation and Technological Change. Blackwell Publishers Inc, UK.

Traxler, G., Byerlee, D., Jain, K.B.L., 1995, The Downsizing of Agricultural Research: Incorporating Scale Economies and Spillovers in the Research Evaluation. Presented at the Meeting of the American Agricultural Economics Association, 6–9 August 1995, Indianapolis, Indiana.

Von Krogh, G., Ichijo, K., Nonaka, I., 2000. Enabling Knowledge Creation, How to Unlock the Mystery of Tacit Knowledge and Release the Power of Innovation. Oxford University Press, NY.

Zenger, T., 1994. Explaining organizational diseconomies of scale in R&D: agency problems and the allocation of engineering talent, ideas and effort by firm size. Management Science 40 (6), 708–729.