Chapter 3

Estimating Impacts of Geographic Information Systems Research: Using Rubbery Scales and Fuzzy Criteria

*Douglas White**

"When you think you know something about a subject, try to put a number on it. If you can, then maybe you know something about the subject. If you cannot then perhaps you should admit to yourself that your knowledge is of a meager and unsatisfactory kind."

Lord Kelvin, 1893

Introduction

Impact assessments of international agricultural research have documented past efforts and guided its future direction. Numerous *ex post* impact assessment studies authenticated early Consultative Group on International Agricultural Research (CGIAR) successes of increased staple grain productivity. For years, such assessments have influenced decisions with regard to the allocation of financial resources. In general, crop research that produced greater economic benefits received larger investments.

The distinction between natural resource management (NRM) and integrated natural resource management (INRM) can be subtle (for more on the evolution of these approaches see Douthwaite et al., 2003; Fujisaka and White, 2003). This chapter does not distinguish between the two terms; INRM is used to describe both these relatively new CGIAR objectives. In addition to increasing agricultural productivity, research objectives added the broader development objectives of (1) alleviating poverty, (2) preserving the environment, (3) spurring economic growth, and (4) facilitating organizational/institutional change. While these objectives were sometimes embedded within earlier research efforts, they have become more explicitly important following the vanguard of modern ecological and social science.

But identifying and measuring the impacts of such an extensive INRM research agenda remains difficult. For example, geographic information systems (GIS) research includes upstream products of knowledge, information, and training that modify decisions and policies, which in turn lead to final impacts. Two major issues confound efforts to assess the impact of GIS research: (1) clearly identifying the multiple cause-and-effect

---

* Agricultural Economist, Land Use Project, Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia.

relationships, and (2) determining appropriate precision of measurement instruments. This chapter's objective is to provide a rapid, comprehensive, and generalized method to evaluate the impact of GIS research. Three objectives of INRM—poverty alleviation, environmental preservation, and economic growth—are used as summary development goals (Reardon and Vosti, 1995). Research outputs affect change with respect to one, two, or all three of these objectives. Two additional measures consider the process by which research influences the development goals: (1) the level of participation and (2) the spatial scale at which research and impacts take place. Impacts are evaluated in a systematic and transparent manner using qualitative criteria. Research outputs of the Land Use Project of the International Center for Tropical Agriculture (CIAT, the Spanish acronym) serve as the case study.

Impact Assessment of INRM Research

Impact assessments are used to improve decision making and resource allocation. They provide an account of past investments, and identify promising and effective investments. Multiple cause-and-effect relationships of INRM research, however, bewilder attempts to assess impact. INRM research is complex in both its approaches and results. Since multiple objectives reflect the needs and expectations of different stakeholders (Izac and Sánchez, 1998), interventions range from relatively tangible germplasm and land management to subtle development processes of increased knowledge and capacity. Functions of organizations, policy, and institutions comprise this latter and larger research domain (Leeuw, 2000).

Demonstrating links between such research outputs and development impacts is difficult. Complications are particularly acute with upstream research products of information and training. A long chain of events is often required where many people adapt and improve a scientific innovation before adoption and impact occurs. In other words, adoption processes are not linear (Douthwaite et al., 2003; Ekboir, 2003; Kuby, 2003). Adopters and researchers work, learn, and affect change together. Furthermore, other concurrent development processes, such as changing government policies and market prices, confound identification and measurement. Hence, causality of research impact upon development process is often tenuous and hard to assess.

Methods

This chapter employs three unconventional methods to comprehensively estimate the impact of INRM research projects. One, multiple evaluation criteria correspond to the numerous objectives and subobjectives of INRM. Two, scientists subjectively assess the impact of their own projects. Three, elicited responses are qualitative measures analyzed using descriptive statistics and correlations between the multiple objectives.

The multiple objectives inherent to INRM research require appropriate recognition and measurement of different impacts. Six summary criteria are used to estimate the impacts of research projects, similar to those used by Campbell et al. (2001) and by Kristjanson and Thornton (2002) of the International Livestock Research Institute (ILRI). Three criteria relate to the development goals of poverty alleviation, environmental preservation, and economic growth; two scalar criteria, level of participation and geographic scale of research impact, take account of INRM processes; one summary criterion estimates the cost of the research (Figure 1). Together, these criteria answer the basic questions of what, when, where, how, who, and how much.

Development goals (*what, when*)
Poverty alleviation (equity)
Environmental sustainability
Economic growth

Scalar adjustments
Geographic (*where*)
Participation (*how, who*)

Research costs (*how much*)

Figure 1. Impact assessment of integrated natural resource management research and questions addressed.

The three development goals form the first half of the summary impact criteria. They encapsulate the key objectives of INRM research—the *what* question. As they appear, however, the three criteria are difficult to understand, and require additional criteria in order to minimize individualistic interpretations. Two of the goals, environmental preservation and poverty alleviation, use a pair subcriteria in order to evaluate the relative importance of the problem and the ability of the research project to affect change. These subcriteria have a temporal aspect to them—the *when* question. Severity refers to the current state of the problem, whereas vulnerability considers future seriousness of the problem being addressed by research. The vulnerability component attempts to address perceptions of fragility or lack or resilience to exogenous shocks. The third goal, economic growth, employs more traditional economic measures. Since the poverty alleviation goal captures aspects of severity and vulnerability, the economic growth criterion is relatively straightforward. A research project is evaluated according to the size of production difference, and what the change represents within overall household income. More sophisticated economic models could be used in order to estimate more accurately the economic benefits, but again they are time consuming and expensive to implement.

Two scalar criteria address the process aspects of INRM research impacts. The first scalar estimates the geographic coverage impact—the *where* question. Given that development processes occur at different organizational levels, INRM research includes higher scales of analysis above the field, plot, and farm. Research may have a tendency to focus on a specific region and have pervasive effects, or it may be wider in scope and influence a lower percentage. To capture this possibility, estimates are made regarding the percentage of people or land area affected at four different scales—community, nation, continent, and globe. In the case of economic impacts, the geographic scalar estimates the rate of adoption per given area at different area scales.

The second scalar refers to the level of participation—the *how* and *who* questions. Research that includes other scientists and development workers is deemed to have higher "buy-in", so impacts have a higher probability of occurring and lasting longer. Beneficiaries of research take on more active roles by determining and implementing the research and development agenda. Such a participatory approach is seen to be more sustainable, following the adage of "teaching a person to fish." As a result, INRM approaches empower many people ranging from farmers and extension agents to policymakers and fellow scientists. Besides improving the potential of individuals, these efforts also build social capital that encourages development processes.

To capture the human and social capital impacts, research outputs are evaluated according to a scale of participation. The scale functions on a cumulative basis. Research that produces scientific journal publications alone has the lowest score. Adding technical reports/Web site/CD-ROM raises the score to the next level. The previous outputs, along with training and the establishment of a user/discussion network, receive a higher score. A demonstrated policy change, at any spatial scale, from community to globe, is the highest level. This scale estimates the level of policymaker empowerment at scales ranging from the farmer, who is a private policymaker/manager, to administrators who may influence policy over much larger spatial areas.

The final criterion addresses research costs—the *how much* question. Research costs are a function of the number of scientists involved, the percentage of time they dedicate to the project, and the number of years the project requires. This estimate also serves as an estimate of the project size.

A case study of intermediate INRM impact: GIS research

Geographic information systems, along with associated spatial analysis, are an example of INRM research that does not lead to direct impacts. Nevertheless, the research does have influence; the challenge is to derive valid estimates of impacts. Scientists of the Land Use Project evaluated their research projects (n = 31) according to the above criteria. The list of research projects is given below. Qualitative measures systematically recorded their subjective assessments. Measures were intentionally imprecise to avoid pseudo-precision. Four categories were employed with scores ranging from 0 to 3; intermediate values were also used (e.g., 2.5). Higher values represent positive, desired traits. Table 1 presents a summary of the criteria, subcriteria, and qualitative scoring scales.

Table 1. Impact assessment criteria of integrated natural resource management research and qualitative scales.

Goal/process criteria	Subcriteria	Qualitative scale descriptors
Economic impact	Production change, percentage of household income	0 = none 1 = low 2 = medium 3 = high
Environmental impact	Severity, vulnerability	0 = negative 1 = neutral 2 = good 3 = excellent
Poverty alleviation impact (equity)	Severity, vulnerability	0 = none 1 = low 2 = medium 3 = high
Geographic coverage (population affected)	Community, national, continental, global	0 = 0%-24% 1 = 25%-49% 2 = 50%-74% 3 = 75%-100%
Participation (level of decentralization)	Scientists, development workers	0 = only scientific journal publications 1 = plus technical reports/Web site/ CD-ROM 2 = plus training/networks 3 = plus policy change
Research cost	No. of scientists, percentage of time, time period (years)	0 = none 1 = low 2 = medium 3 = high

Research projects of the CIAT Land Use Project

Accessibility and spatial interaction analysis
Basic needs index for Central America
Cassava resilience on hillsides
Climate database
Consortium of Spatial Information
Decision support system (DSS) for agricultural projects and land use
DSS of Andean infrastructure
Ecoregional research network
FloraMap
Food insecurity mapping (Ecuador)
Food security and poverty mapping
Genotype selection in participatory bean experiment
High spatial resolution imagery
Landslide prediction
Land use change (Nicaragua)
Land use planning training (Ecuador)
Local and scientist views of NRM
Maize and climate change
MarkSim
Measure/model forest biodiversity
Participatory 3-D mapping
Remote sensing for planning
Role of local knowledge in NRM
Rural sustainability indicators
Socio-spatial decisions of forages
Soil macrofauna at catchment scale
Spatial interactions of dairy markets
Targeted wild relatives conservation
Tropical precision agriculture
Whitefly and climate change
Wild beans and climate change

All subcriteria except those of geographic scale are equally weighted. Since a central objective of the CGIAR is to produce international public goods, research that affects change at larger geographic scales receives greater weight. Amongst the four categories (community, national, continental, and global), the two lower scales use a multiplier of 0.2, while the two higher scales use a multiplier of 0.3, thereby summing to one.

Scientists scored their research projects within an electronic spreadsheet. Estimates were not made in isolation; scientists compared their evaluation scores with those of other projects. Cells of the spreadsheet acquired darker hues as scores increased in order to facilitate rapid visual recognition of the score and comparison with other project assessments. The survey instrument was administered with the author present to clarify questions.

Systematic inquiry of the development goal, processes, and costs enables the examination of various hypotheses:

H_o: Survey results of GIS research projects are homogenous.
Scientists will be unwilling to distinguish the potential impacts of their research outputs.

H_o: Perceived impacts of GIS research are equal with respect to three development goals.
Research projects are multi-objective; investments demonstrate a balance amongst the goals.

H_o: Research at higher spatial scales is inversely related to decentralized research approaches.
Participatory approaches typically occur at local levels. GIS research and analysis at higher scales, as with policymakers, is rarely collaborative.

H_o: Higher cost research is more decentralized.
Participatory research processes require more time to coordinate efforts and have expensive travel costs.

H_o: Research to alleviate poverty spurs economic growth.

H_o: Poverty alleviation research focuses on site-specific regions.
The issues of poverty are highly contextual requiring in-depth analysis of geographic regions.

Results

Analysis of the qualitative data provides numerous insights into how GIS scientists view the influence of their research. The qualitative data enabled rapid and systematic examination of general interrelationships between development goals, processes, and costs. Quantitative summary statistics (e.g., mean, standard deviation, and correlation coefficients) were used to analyze the elicited scores. In comparison to quantitative analyses, the potential of qualitative analyses to make detailed inferences has many limitations. The qualitative measures employed do not use a common metric; therefore, results amongst the different measured criteria are not directly comparable.

Scientists estimated modest impacts of their research on average (Figure 2). GIS research was seen to have similar qualitative effects (~1.6) on economic growth and environmental preservation. Since the two qualitative scales differ, these translate into medium-low impact on economic growth, and between neutral and good impact on the environment. Research toward the equity goal had stronger perceived impact (1.9). This result reflects the poverty alleviation strategy and tactics of the GIS project.

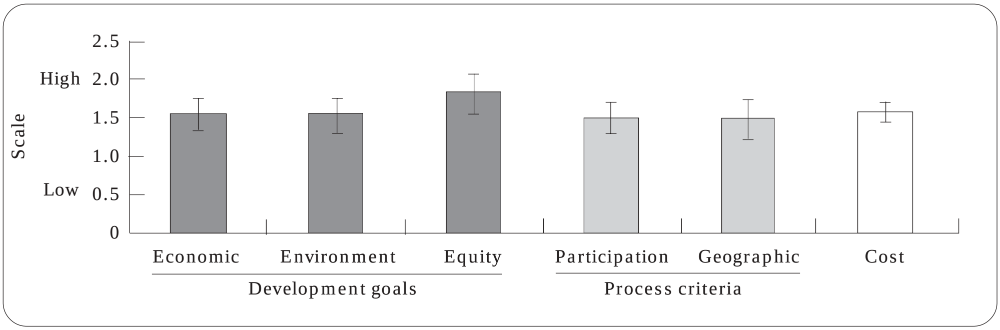

Figure 2. *Ex-ante* impact assessment estimates of Land Use Project (mean, standard deviation).

The standard deviations about the mean of the elicited responses were similar, about 0.46 for the economic, environment, and participation criteria. These somewhat large standard deviations imply that researchers were able to distinguish different levels of impact of the projects and rate them accordingly. The process criteria of INRM also received medium-low ratings along with the cost index. Again, all comparisons between the indices must be made with care; elicitation of responses was accomplished by evaluating a research project per criteria. No assessments regarding the relative importance of the criteria were conducted. These summary results are more a demonstration of the behavior of the indices than a comparison between the distinct criteria.

Correlation coefficients examine general tendencies of the qualitative data and produce logically consistent results (Table 2). The correlation coefficients compare the entire group of projects with respect to the development goals, processes, and costs. Some results were anticipated; others were not. An example of an expected result is that research addressing economic growth is highly correlated (0.62) with poverty alleviation (equity). The criteria appear to have much thematic overlap. Also, impacts of the development goals are highly correlated with research costs, ranging from 0.40 with economic development to 0.67 of environmental preservation. Surprisingly, however, more participatory approaches are only slightly positively correlated with research costs, 0.25. At the risk of pseudo-precise results, correlations greater than 0.37 are statistically significant at $\alpha = 0.05$.

Table 2. Correlation coefficients of impacts of geographic information systems research.

	Environment	Equity	Participation	Geographic	Cost
Economic	-0.11	0.62*	-0.06	0.27	0.40*
Environment		-0.08	0.36	0.20	0.67*
Equity			-0.22	-0.23	0.59*
Participation				-0.30	0.25
Geographic					0.42*

* statistically significant ($r > 0$, $\alpha = 0.05$).

The levels of research participation are positively and negatively correlated with the different development goals. Participatory methods are positively correlated with the environmental research, but negatively correlated with themes of economic growth and equity. This result may be due to the management requirements of natural resources by local people, whereby participatory approaches are more effective. A negative correlation between increased participation and larger geographic scales (–0.30) appears to support this result. Also, research at the community level reveals a tendency of employing a more participatory approach (correlation coefficient = 0.38).

Of the six criteria examined, geographic scale contains the most subcriteria. A more detailed analysis of responses reveals that scientists perceive that their research has more pervasive impact at smaller scales (Figure 3). Average assessments of impact, equally weighted, range from medium (1.9) at the community level to just above low (1.1) for global. The

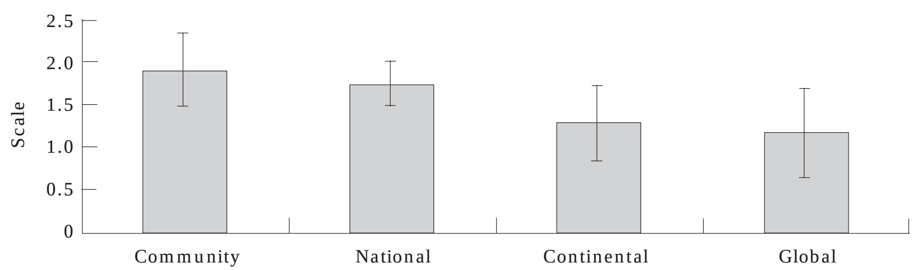

Figure 3. Perceived impacts of research per geographic scale (mean, standard deviation).

variation about the mean also changes according to scale. From national to global scales, standard deviation becomes larger as scale increases. The standard deviation of GIS research impacts at the community level was relatively high.

Correlation coefficients of the geographic subcriteria are both negative and positive (Table 3). Community level research is negatively correlated with all higher scales, ranging from –0.13 at the national level to –0.59 at the continental level. This implies that research that has greater perceived impact for specific communities is not easily generalized. Research impacts at higher scales are positively correlated. This could mean that, once beyond the community level, research impacts are generally applicable and scales have less distinction and fewer implications. The high positive correlation between global and continental research supports such an inference.

Table 3. Correlation coefficients of impacts of geographic information systems research at geographic scales.

	National	Continental	Global
Community	-0.13	-0.59*	-0.36
National		0.45*	0.29
Continental			0.70*

* statistically significant ($r > 0$, $\alpha = 0.05$).

Returning to the hypotheses posed in the previous section, many of them were founded. Scientists were willing and able to distinguish the potential impact of their research outputs. Variability in responses was reflected by the standard deviations about the means. Scientists also assessed different levels of impact to the three development goals. The overall mean of the three goals by project was 1.6, with a standard deviation of 0.4. This is the average of the means, which is not equal to the mean of the averages (1.7) as depicted in Figure 2.

Participatory GIS research is more costly. Time required to coordinate research with others is likely to be longer than a scientific publication strategy. Travel costs are also likely to increase when more people are involved. With respect to specific development goals, some analysis outcomes are expected. Results fail to reject the hypothesis that GIS research impacting economic growth also alleviates poverty. The two impacts of research projects are highly correlated (0.62). In contrast, GIS research that addresses poverty alleviation does not necessarily occur at a community scale. Projects demonstrate a nearly negligible positive correlation (0.07).

Research benefits of GIS (i.e., impacts) tend to increase as costs increase (Figures 4 and 5). Both cost and benefit estimates demonstrate sufficient dispersion, supporting the inference that scientists could distinguish their work using qualitative measures. Vertical groupings of results are an artifact of the categorical nature of the cost estimates and equal weighting of the subcriteria. Scientists tended to respond using integers and in-between half values (e.g., 2.5). This could signify that insufficient detail was provided with the subcriteria and associated scales, thereby causing scientists to respond with broad estimates.

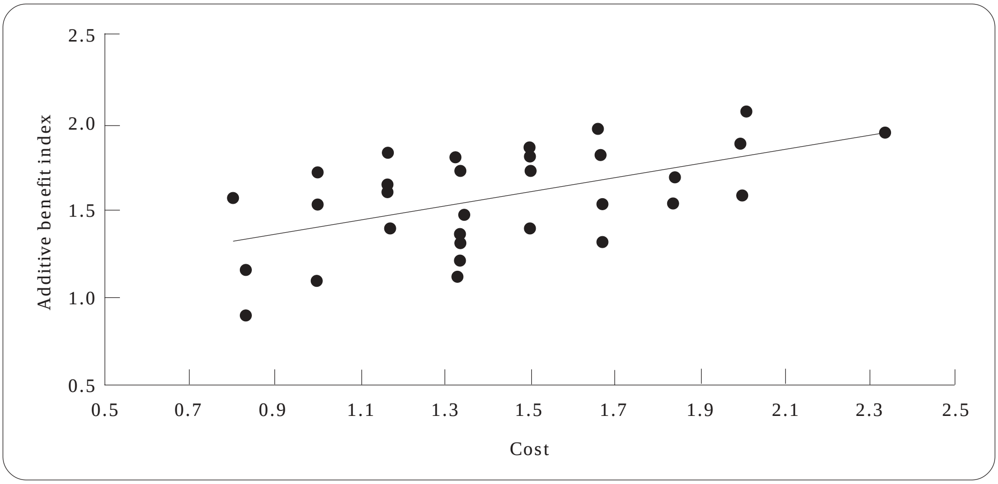

Figure 4. Benefits (additive scales) versus costs of geographic information systems projects.

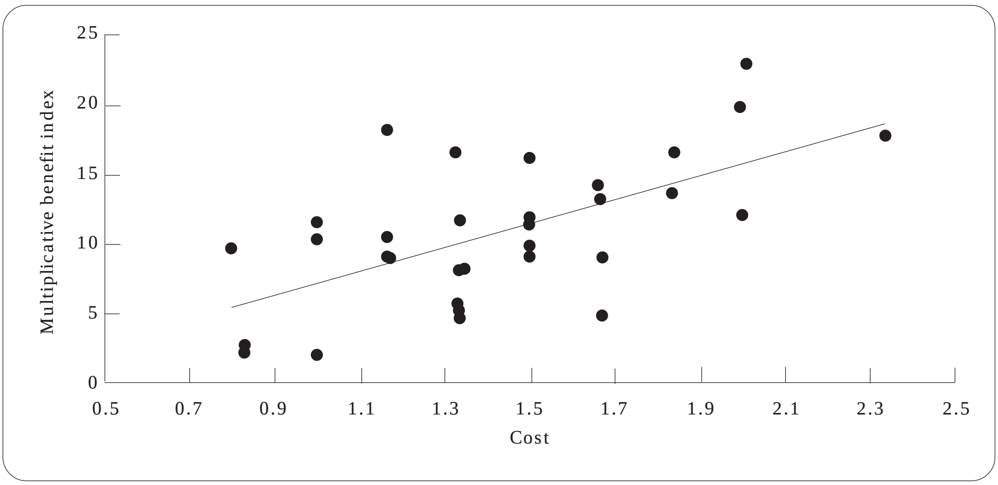

Figure 5. Benefits (multiplicative scales) versus costs of geographic information systems projects.

Summary analyses using two different associations amongst the criteria reveal different results. Figure 4 is based upon a simple additive association where total score is the sum of the five criteria. All the criteria are weighted equally. In Figure 5, the three development goal impacts were scaled by the level of participation and geographic coverage. The process criteria were used as multipliers. As a result, the summary scores of many of the projects change, as can be seen by the different estimate positions within the vertical groupings.

Discussion

The use of unconventional methods to estimate the impact of INRM research projects raises more questions than it answers, especially with regard to (1) goals and their definitions, (2) multiple interpretations and measurement, and (3) their relation and analysis.

One, the comprehensive nature of INRM research requires that impact assessment include multiple evaluation criteria that correspond to numerous objectives and subobjectives. But achieving accurate measures against criteria is another matter. Given the complexity of research projects and their impact context, the use of precise measures would be invalid. The goals of INRM research are subjective concepts that are not only ill defined, but also distortable by emotion or personal bias. Despite estimation challenges, many scientific disciplines attempt to objectively measure subjective phenomena. Psychologists, for example, estimate intelligence and personality traits (Dalkey and Rourke, 1971). Such characteristics are imprecisely defined, and thus open to interpretation. Similarly, economists use survey instruments to measure subjective characteristics, for instance consumer preferences. Although these types of estimates are not considered to be highly exact, they provide a basis with which to analyze difficult-to-define subject matter. The initial broad tendencies can be identified, contrasted, and further explored.

Two, eliciting expert opinion is one manner with which to estimate research impacts. Personal biases and preferences, however, can affect responses. Overstating research impact is a tempting strategic behavior to satisfy desires of professional advancement or personal ego. Although personal subjective judgments remain within an evaluation, the transparent peer-review evaluation process minimizes such potential behavior by providing a checks-and-balances system.

Many concepts are ill defined because of multiple themes embedded within them. Poverty, for example, is a well-known concept, but difficult to fully characterize. Besides a World Bank definition of income being less than US$1 per day, other aspects of the condition require recognition, such as empowerment, opportunity, and nourishment. Thus, the use of subcriteria that represent aspects of the larger concept facilitates more general understandings and reduces personal interpretations.

Three, qualitative measures limit the ability to conduct rigorous quantitative analysis. Partly as a result of unclear goal and criteria definitions, this study relies upon direct comparison between projects in order to estimate impacts. Since the categorical evaluation scales do not always provide consistent interval measures (Scheibe et al., 1975), the associated numerical values need to be analyzed with caution. For example, the concepts and criteria of the poverty alleviation and economic growth development goals appear to overlap. This may lead to problems of double counting, which in turn may skew summary results away from the environmental preservation goal. Similarly, the process criteria of GIS research, participation and geographic scale, could benefit from further refinement. Their relationship to the development goals, whether additive or multiplicative, also requires discussion.

The weights of the indices are subjectively determined. Analysis results directly depend upon the weights, since they determine the relative importance of the criteria. The ILRI study, for example, established their relative values via expert opinion. Yet other views of diverse INRM stakeholders are also important to consider (Kelley et al., 1995). Future research could contrast the preferences and priorities of stakeholder groups with analytical hierarchy process (Saaty, 1995) or Delphi (Turrof, 1970) methods.

To describe research to development processes, causal pathways are often used to explicitly document the intermediate links between the final impacts of research outputs (Gottret and White, 2001; Douthwaite et al., 2003). Pathways help establish a plausibility of impact by explaining the context and identifying conditions or concurrent interventions that are required in order for impacts to occur. Three points along the path are distinguished: (1) outputs, immediate products of a project after using the given inputs, (2) outcomes, consequences of the outputs, and (3) impacts, the broader and longer-term goals. Scientist responsibility and control over specific activities declines as one moves along the pathway from a research output to a development impact (Smutylo, 2001). Causal pathways, however, are difficult to compare since no summary measure are developed. The participation criteria used in this chapter attempt to estimate the strength and magnitude of the pathway links. Indicators of participation could include more detailed assessment of processes such as those of Biggs (1989) and Lilja and Ashby (1999).

Questions of analytical rigor

Quantitative economic impact assessments of Green Revolution crop improvement research established a high standard for broader INRM impact assessment approaches to meet. A single monetary value describing research benefits has indisputable appeal when making decisions. Such an estimate is easy to comprehend and compare with other research efforts.

Impact assessments come in many forms and differ in analytical rigor. On one side of a continuum representing different levels of rigor are quantitative impact assessments (Figure 6). Systematic and mathematically sophisticated methods provide objective estimates of research impacts. Most of these concern economic impacts (Pardey et al., 1991; Alston et al., 1995; 2000). Econometric models are often used to estimate not only the overall magnitude of benefits, but also how these benefits are distributed, such as toward the poor (Binswanger, 1980; Ravillion, 2001). Use of a common metric, a monetary measure such as US dollars, facilitates comparison amongst different studies.

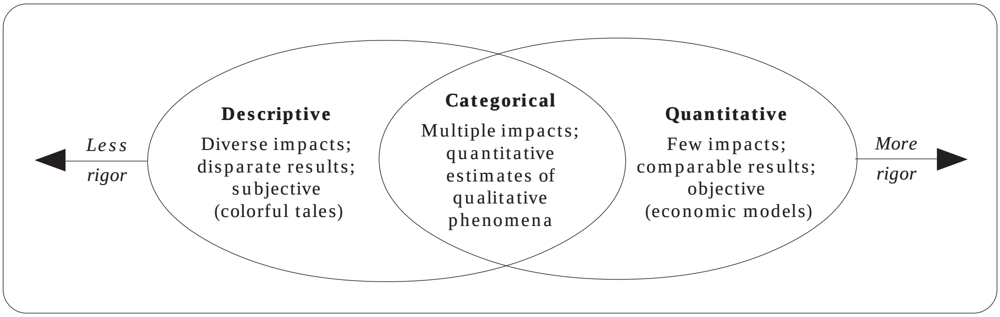

Figure 6. An impact assessment method's continuum, representing different levels of analytical rigor.

On the other side of the continuum are descriptive impact assessments. Despite their rather informal and subjective methods, they are often persuasive. Anecdotes of research success include tales of improved farm earnings, increased farmer participation in research processes, or sustainable management of resources. Such human-interest stories can be effective in conveying to listeners and readers that impact has been achieved. The relevance and potential impact of individual successes can, in theory, be scaled out to larger populations and geographic areas by posing plausible arguments regarding others who face similar conditions and challenges.

These two extremes of impact assessment, quantitative and descriptive, tend to measure different types of impact. Rigorous studies typically focus on research outputs that address only one or two of the development goals. Such studies usually concern private or on-farm economic benefits or the public economic benefits of research for a specific commodity (e.g., rice [*Oryza sativa* L.], maize [*Zea mays* L.]) over a larger geographic region. In contrast, descriptive studies are used to explain the benefits of multi-objective INRM-type research, especially improvements in development processes that are difficult to measure. Since many actors and scales are involved, these benefits are often public in addition to private in nature. Rarely are these studies conducted over large geographic areas, but focus on groups of farmers or specific communities (Schioler, 1998; 2002).

Between these extremes appear qualitative impact assessment approaches. These studies often employ both non-economic quantitative and qualitative measures to estimate diverse impacts of research (see Horton et al., 1993). Indicators and indices summarize before and after conditions to estimate impacts. Participatory monitoring and evaluation fits into this realm (Guijt, 1998). While this scoring approach can address a broad research and development agenda, it tends to be site specific. Increased local participation highlights local concerns, and thereby reduces the ability to compare results with other impact assessments (Gottret and White, 2001).

Conclusion

The methodological approach used in this chapter appears to both conflict and concur with those of the recent literature. In many cases, evaluators should seek to "establish plausible links" between research investments and development impacts rather than to "prove causation" or "measure impacts" of research on summary development goals (EIARD, 2003).

No matter how well intended or well developed evaluative activities are, they can and probably will have unintended and undesired side effects, thereby jeopardizing effectiveness and performance. One way around such an uncomfortable result is to perceive evaluation as providing a learning function that facilitates knowledge building in the collaborative development contexts. More then ever before, larger numbers of different stakeholders are involved in evaluation and impact assessments (Leeuw, 2000; Horton and Mackay, 2003).

Raising questions is perhaps one of the latent objectives of this chapter. Despite analyzing concepts with fuzzy generalized criteria and rubbery scales, discussion of how to measure the impacts of research projects spurs further analysis of how to upstream research outputs more effectively.

References

Alston, J.M.; Norton, G.W.; Pardey, P.G. 1995. Science under scarcity: Principles and practice for agricultural research evaluation and priority setting. Cornell University Press, Ithaca, NY, USA. 585 p.

Alston, J.M.; Chan-Kang, C.; Marra, M.C.; Pardy, P.G.; Wyatt, T.J. 2000. A meta-analysis of rates of return to agricultural research and development: *Ex pede herculem*? International Food Policy Research Institute (IFPRI) Research Report no. 113, Washington DC, USA. 148 p.

Biggs, S.D. 1989. Resource-poor farmer participation in research: A synthesis of experiences from nine national agricultural research systems. *In*: On-farm (client-oriented) research (OFCOR) Comparative Study Paper 3, International Service for National Agricultural Research (ISNAR), NL. p. 3-37.

Binswanger, H.P. 1980. Income distribution effects of technical change: Some analytical issues. South East Asian Econ. Rev. 1(3):179-218.

Campbell, B.; Sayer, J.A.; Frost, P.; Vermeulen, S.; Ruiz Pirez, M.; Cunningham, A.; Prabhu, R. 2001. Assessing the performance of natural resource systems. Conserv. Ecol. 5(2):22. [online] Available in: http://www.consecol.org/vol5/iss2/art22/index.html

Dalkey, N.S.; Rourke, D.L. 1971. Experimental assessment of Delphi procedures with group value judgment. Rand Corporation, R0612-ARPA. 58 p.

Douthwaite, B.; Kuby, T.; van de Fliert, E.; Schulz, S. 2003. Impact pathway evaluation: An approach for achieving and attributing impact in complex systems. Agric. Syst. 78:243-265.

EIARD (European Initiative for Agricultural Research for Development). 2003. Impact assessment and evaluation in agricultural research for development. Agric. Syst. 78:329-336.

Ekboir, J. 2003. Why impact analysis should not be used for research evaluation and what the alternatives are. Agric. Syst. 78:166-184.

Fujisaka, S.; White, D.S. 2003. *Ex post* methods to measure natural resource management research impacts. Paper commissioned by the Center Directors Committee of the Integrated Natural Resource Management (INRM) Task Force. Centro Internacional de Agricultura Tropical (CIAT), Cali, CO. 26 p.

Gottret, M.V.; White, D.S. 2001. Assessing the impact of integrated natural resource management: Challenges and experiences. Conserv. Ecol. 5(2):17. [online] Available in: http://www.consecol.org/vol5/iss2/art17

Guijt, I. 1998. Participatory monitoring and impact assessment of sustainable agricultural initiatives: An introduction to key elements. SARL Discussion Paper no. 1, International Institute for Environment and Development (IIED), London, GB. 112 p.

Horton, D.; Mackay, R. 2003. Using evaluation to enhance institutional learning and change: Recent experiences with agricultural research and development. Agric. Syst. 78:127-142.

Horton, D.; Ballantyne, P.; Peterson, W.; Uribe, B.; Gapasin, D.; Sheridan, K. (eds.). 1993. Monitoring and evaluating agricultural research: A sourcebook. CAB International, Wallingford, GB. 219 p.

Izac, A.M.; Sánchez, P.A. 1998. Towards a natural resource management paradigm for international agriculture: Example of agroforestry research. International Center for Research in Agroforestry (ICRAF), Nairobi, KE. 26 p.

Kelley, T.G.; Ryan, J.G.; Patel, B.K. 1995. Applied participatory priority setting in international agricultural research: Making trade-offs transparent and explicit. Agric. Syst. 49:177-216.

Kristjanson, P.M.; Thornton, P.K. 2002. Herding cats: Is impact assessment the ultimate exercise in futility? Paper presented at the International Conference on the Impacts of Agricultural Research and Development, "Why has impact assessment research not made more of a difference?", 4-7 February, San José, CR.

Kuby, T. 2003. Innovation is a social process: What does this mean for impact assessment in agricultural research? *In*: Mathur, S.; Pachico, D. (eds.). Agricultural research and poverty reduction: Some issues and evidence. Economics and Impact Series no. 2, Centro Internacional de Agricultura Tropical (CIAT), Cali, CO. p. 59-70.

Leeuw, F.L. 2000. Program evaluation and social and institutional impact assessment. Workshop organized by the Standing Panel on Impact Assessment (SPIA) of the Technical Advisory Committee (TAC) of the Consultative Group on International Agricultural Research (CGIAR) "The future of impact assessment in the CGIAR: Needs, constraints and options", 3-5 May, Food and Agriculture Organization (FAO), Rome, IT.

Lilja, N.; Ashby, J. 1999. Types of participatory research based on locus of decision making. Consultative Group on International Agricultural Research (CGIAR) Systemwide Program on Participatory Research and Gender Analysis (PRGA), Centro Internacional de Agricultura Tropical (CIAT), Cali, CO. 8 p. (PRGA Working Document no. 6)

Pardey, P.G.; Roseboom, J.; Anderson, J.R. 1991. Agricultural research policy: International quantitative perspectives. Cambridge University Press, Cambridge, MA, USA. 462 p.

Ravillion, M. 2001. The mystery of vanishing benefits: An introduction to impact evaluation. World Bank Econ. Rev. 15(1):115-140.

Reardon, T.; Vosti, S.A. 1995. Links between rural poverty and the environment in developing countries: Asset categories and investment poverty. World Dev. 23:1495-1506.

Saaty, T.L. 1995. Decision making with the analytical hierarchy process. Int. J. Manage. Syst. 10(1-3):219-258.

Scheibe, M.; Skutsch, M.; Schofer, J. 1975. Experiments in Delphi methodology. *In*: Turoff, M.; Linstone, H. (eds.). The Delphi method: Techniques and applications. Addison-Wesley, Reading, MA, USA. Available in: http://www.is.njit.edu/pubs/delphibook/ch4c.html

Schioler, E. 1998. Good news from Africa: Farmers, agricultural research, and food in the pantry. International Food Policy Research Institute (IFPRI), Washington D.C., USA. 72 p.

Schioler, E. 2002. From the rural heart of Latin America: Farmers, agricultural research, and livelihoods. Future Harvest, Washington D.C., USA. 80 p.

Smutylo, T. 2001. Crouching impact, hidden attribution: Overcoming threats to learning in development programs. Draft Learning Methodology Paper presented at the Block Island workshop on across-portfolio learning, 22-24 May, Evaluation Unit, International Development Research Centre (IDRC), Ottawa, CA.

Turrof, M. 1970. The design of a policy Delphi. J. Technol. Forecast. Soc. Change (2)2.