gtz
Deutsche Gesellschaft für
Technische Zusammenarbeit (GTZ) GmbH

ESTABLISHING PLAUSIBILITY IN IMPACT ASSESSMENT

By the Workgroup on Assessing the Impact of Agricultural Research in Development:

Henning Baur, Michael Bosch, Stephan Krall, Thomas Kuby, Alison Lobb-Rabe, Paul-Theodor Schütz, and Andreas Springer-Heinze

Eschborn

February 2003

Abstract: On the basis of an IITA impact study, this paper demonstrates the importance of establishing plausible linkages between development cooperation interventions and the relevant changes observed on the ground when assessing impact. It also proposes standards for establishing plausibility in this context. Experience has shown that impact evaluators often try to prove or quantify impact on a highly-aggregated level. However, this paper contends that in most cases, attribution gaps caused by the existence of too many other significant factors make it impossible to isolate the effects of a single development intervention. The authors maintain that, for most development interventions, to require impact evaluators to establish more than plausible impact relationships would force them to gloss over much information and over-interpret the data.

IMPACT ASSESSMENT AND PLAUSIBILITY

The value of impact assessment to development cooperation has been increasingly recognized over the past several years, as evidenced by the growing body of literature and reports in this area. Impact assessment supports development interventions in a number of ways, the most important of which are:

- *learning* about more and less successful approaches to development and poverty reduction
- *steering* projects, programs, and strategies within their given, dynamic settings to maximize effectiveness and sustainability
- *improving accountability* for investments in development cooperation by trying to ensure that they truly effect changes in the lives of people, especially the poor.

Obviously, a superficial impact report that sets an intervention in as positive a light as possible cannot serve these functions. On the other hand, it usually is not possible to determine a scientifically sound, discontinuity-free, cause-and-effect relationship between an isolated development measure or strategy and the changes observed on the ground. We maintain that the reasonable compromise is **to establish plausible links** between the development efforts and the observed changes. We find this compromise to be reasonable because 1) it is possible to achieve using an appropriate amount of resources, and 2) it delivers meaningful information for the purposes of learning, intervention steering, and accountability.

Indeed, this paper proposes that constructing and examining the plausible links to impact is more than a reasonable compromise, but is in fact the central task of impact assessment. To require more than this usually places the impact evaluator in an untenable position. The expectation that a study prove an unbroken and calculable impact chain from intervention or program to changes in people's lives forces the investigators to gloss over reality: that

innumerable factors play roles that are often non-quantifiable, but even more significant than the intervention.

Drawing on Impact Assessment Experience to Derive Guidelines for Establishing Plausibility

The Workgroup on Assessing the Impact of Agricultural Research came together in order to formulate working standards for establishing plausibility in the assessment of impact of development cooperation efforts. We did this by examining a practical example in the area of agricultural research to see what it could reveal about the construction of plausible impact relationships. The example is the International Institute for Tropical Agriculture (IITA) study "Social impact of soybean in Nigeria's southern Guinea savanna" by P.C. Sanginga, A.A. Adesina, V.M. Manyong, O. Otite, and K.E. Dashiell.[1]

The study lays a good basis for a discussion of establishing plausible links to impact because it considers the effects of a development intervention along a prolonged impact chain. In assessing the impact of new varieties of soybean and new production and utilization technologies, the IITA study considers not only such aspects as crop yields and even economic effects for adopting farmers, but also social impact, such as changes in community and household well-being.

"Social impact of soybean in Nigeria's southern Guinea savanna" begins with a brief history of the development and introduction of improved soybean varieties and utilization technologies in Nigeria. It then describes the methodology used for the assessment of impact, namely the social impact assessment (SIA) framework, based on a technology diagnostic-diffusion-adoption-impact continuum. In other words, the point of entrance for the study was "...accurate diagnosis and understanding of the social organization of productive activities" (p. 5); followed by the "...analysis of the technology diffusion and adoption process." (Ibid.) The final step of the framework is the impact analysis, or "...finding out how far the introduction of an improved technology has been successful in meeting socioeconomic objectives, and how well improved agricultural technologies have satisfied the needs and priorities of households and other units in the target population." (P. 6). Other aspects of the methodology (study zone and population, data collection and analysis) are also presented.

[1] The study is available in its entirety from the IITA website at: www.iita.org/info/impact/soybean.pdf .

The study then analyzes and discusses gender roles in soybean production, the diffusion process of improved soybean varieties, rates and determinants of adoption, several types of impact (effects on household income, farmers' welfare, household food security, children's nutritional status, use and allocation of resources), and impact at the levels of the farm and the community. The study finds that improved soybean varieties were widely adopted by farmers in the study region, and that soybean production had a positive impact on farmers' income for both men and women. Other positive impacts found include: improved household food security and welfare, improved children's nutritional status, and greater investment in human capital.

The IITA impact study should inspire others to examine the social impact of their own interventions. Our workgroup is indebted to the study's research team for providing real-life impact assessment experience that we could critically peruse in order to recognize and test standards for establishing plausibility in impact assessment.

THE PROPOSED STANDARDS FOR ESTABLISHING PLAUSIBILITY IN IMPACT ASSESSMENT

This paper presents the standards we came up with. They bear much similarity with the standards set for qualitative research and evaluation, in general (see the attachments). It is not our intention that the proposed standards be applied rigidly and religiously, but rather that they serve as a frame of reference for impact assessment. We hope that they make the task of assessing impact easier and the findings of impact assessment studies more useful to a larger audience.

Only in the most seldom cases can the impact of a development intervention or strategy be *proven* in the sense that its effects can be isolated and measured. This means that the impact evaluator's job is to demonstrate plausible relationships between the intervention activities and the observed changes in the situation experienced by the target groups and beneficiaries. The establishment of plausibility relies primarily on well-founded argumentation rather than the presentation of airtight proofs. The elements we find to be essential for establishing the plausibility of impact are explicit statements or discussion of:

1. the source of the impact being investigated (what or whose impact is the focus of the study)
2. the model or concept of impact used by the impact evaluators and how it applies to the case at hand

3. the objectives of the impact assessment, as well as its limitations
4. the specific theory of action on which the intervention or strategy has been based (or at least that the researchers are able to construct)
5. the impact hypotheses that were tested by the impact assessment
6. other factors (apart from the intervention) that could have affected the observed changes and alternative impact hypotheses
7. other informed opinions that support and contest the study findings.

In the following pages, we present each of these standards and how they underpin plausibility, drawing on the IITA study for illustration.

1. The source of the impact being investigated (what's or whose impact is the focus of the study). We argue in this paper for constructing plausible links between an impact "generator" and the observed impact; so obviously the establishment of plausibility relies, first of all, on identifying what or who that generator is. The identification and description of the impact source being investigated thus is the foundation for establishing plausibility.

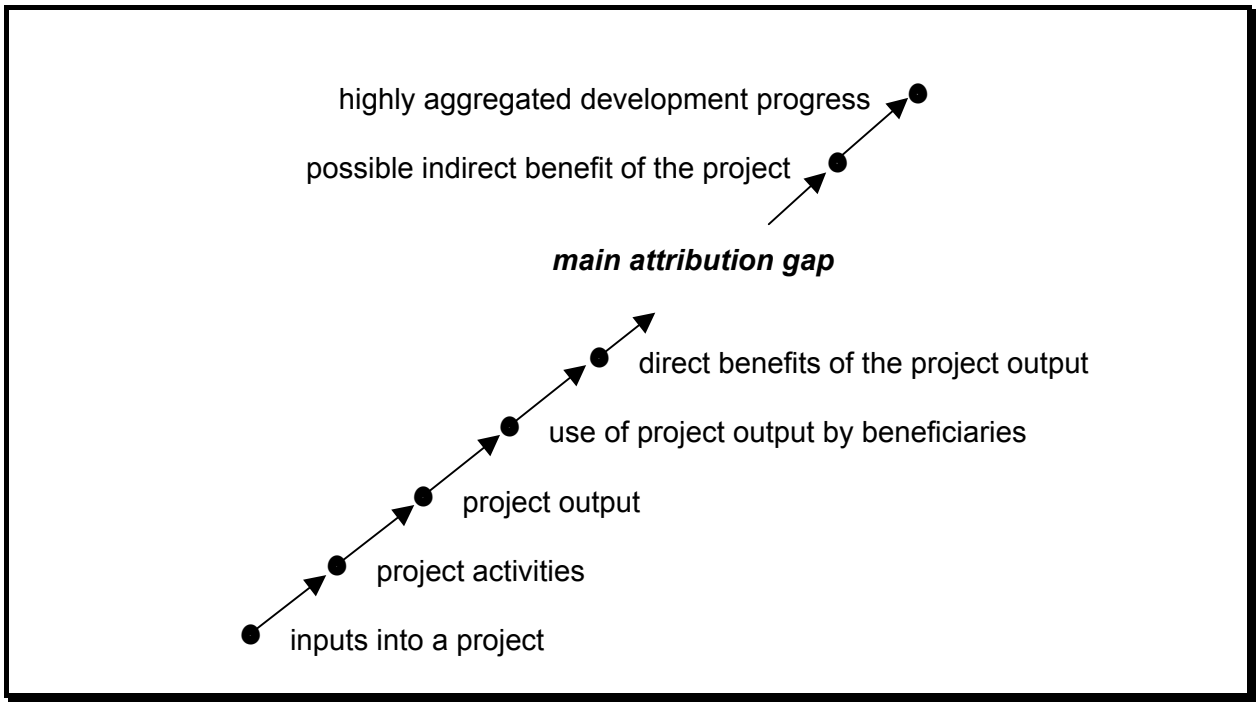
The source of the impact could be the inputs, activities and outputs of a development intervention. It might be stipulated to be a particular program approach or a certain agency. By naming the source of the impact being investigated, the study fixes the point of reference for the impact observations. A list and brief description the activities and processes through which the source ended up generating or influencing the observed changes forges the first links in the chain of plausible impact. This also gives the reader the opportunity to construct his or her own intellectual tests for the findings (such as the degree of magnitude of impact that could be expected).

The IITA study is not always clear on what source of impact it investigated. It sometimes seems that IITA support of the development of the soybean sector in Nigeria was the source of the impact. But page two of the paper states that it concentrates on the impact of "... the introduction of soybean production and utilization technologies to Nigeria." Then we find the statement that, regarding adoption, the study focuses on the five most popular varieties of soybean (p. 4), while the discussion on diffusion is said to concentrate on two varieties (p. 10). It could be a perfectly good approach to look at the impact of a sector as a whole, at the diffusion and adoption of selected varieties, and at the contribution to impact made by a single agency. However, the IITA study does not set up a clear hierarchy or sketch of such

relationships. We therefore found the study at times confusing in that we were not sure what source of impact was being discussed in different passages. Also, although the production and processing technologies are mentioned in the abstract throughout the study, the only one even briefly described is the small plot adoption technique (SPAT) for introducing new varieties (ibid.). A few descriptive sentences about the other production and processing technologies involved would have increased our understanding of how the impact was generated.

2. The model or concept of impact used by the researchers and how it applies to the case at hand. All researchers who investigate impact have an abstract, conceptual picture in their mind of what impact is and how it works – a "map" of how a source of impact links to observed changes in the target population. By making this picture explicit and superimposing the intervention or strategy onto it, we not only help our readers to understand the internal logic of our study, but also check the completeness of our treatment of the case.

We illustrate this with the GTZ Impact Model, below, which is a simplified representation of the complex interrelationships of the many factors involved in creating any impact.

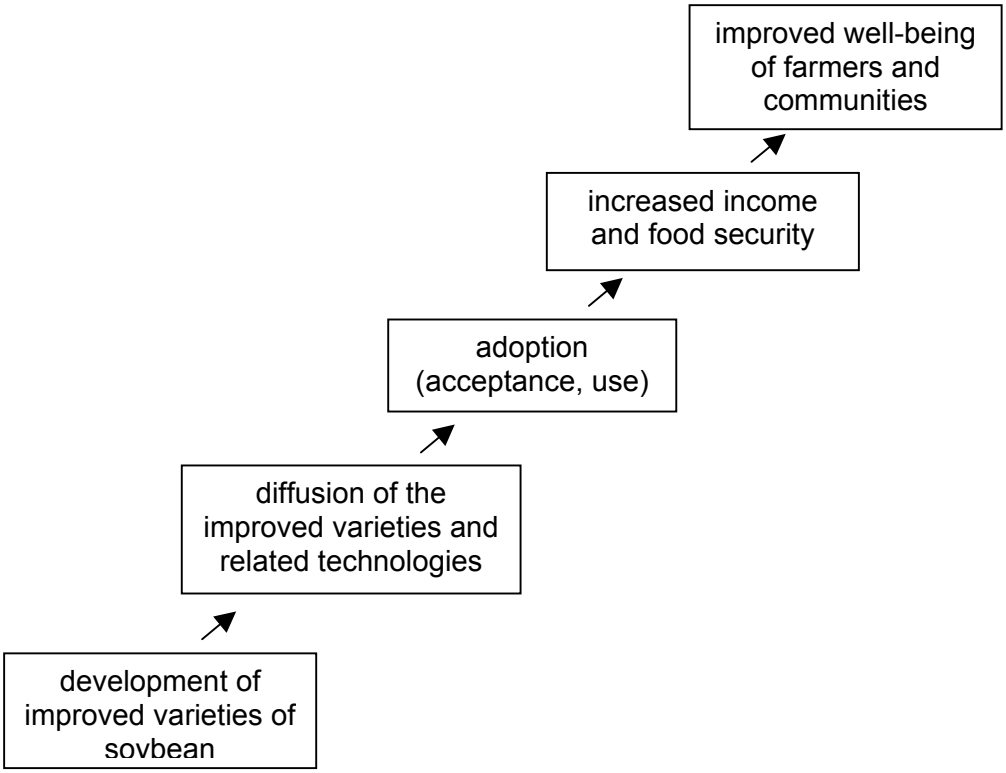


This generic model demonstrates an understanding of impact as "flowing out of", or set in motion by, intervention inputs, activities and outputs. Creating output is not enough to achieve the desired impact, but the output (or services) provided by the project must be used by people. "Use of project output" is where impact begins in this model. Following further along this impact continuum are the direct benefits of the project that the beneficiaries realize through their use of the output/services. Not very far after that, the continuum is broken by an "attribution gap",[2] beyond which it is not possible to attribute observations reliably to the intervention. That is, so many other factors influence these highly-aggregated changes and developments that one cannot isolate the exact role or magnitude of influence the intervention has had at the levels of indirect benefits and "development progress". Hence the need for establishing plausibility in order to span the attribution gap.

The IITA study does not present a conceptual framework of impact. It does present the diagnostic-diffusion-adoption-impact continuum of the SIA framework it used for investigating impact we mentioned earlier. But instead of being illustrating how impact "works", the continuum essentially outlines how an understanding of impact must progress. In presenting the SIA framework, the study also defines its impact analysis task to be examining how and to what the degree the introduction of the new technology has met (unspecified) socioeconomic objectives as well as the needs and priorities of the target groups (p. 6). Then the study lists the many impact indicators that the research team found to be relevant.

However, the study does not give us a picture of how these impact areas and indicators conceptually fit together. One could, for instance, draw on the technology diagnostic-diffusion-adoption-impact continuum to construct the impact "pathway" from the development of new soybean varieties/technologies to the socioeconomic changes experienced by the Nigerian farmers and communities. According to our reading of the study, the picture would look something like this:

[2] Actually, attribution becomes increasingly difficult along the continuum beyond "use of project output" – there is not just a singly attribution gap.



A simple illustration or brief description of the impact concepts operating in the study and how the case at hand fits into them would help to structure the argumentation that establishes plausibility, as well as foster the reader's understanding of it.

3. The objectives and limitations of the impact assessment. In order for the reader to know what to expect from an impact study, the researchers must state what they attempted to do (what purpose the study serves), and the (to a large degree practical) limitations with which they had to work. Explicitly stating the objective and limitations of a study is part of the "ethic of discourse". By being forthcoming about the aims and weaknesses of the study, the author puts the reader on as nearly equal footing as possible in the (albeit unavoidably one-sided) discussion of impact.

Honesty, openness and transparency with the study objectives and limitations also play a role in establishing plausibility. For one, they set the tone of the study for accuracy and completeness of data presentation and analysis. The attribution gap is one limitation that every impact study will encounter. Impact evaluators must be forthright about the existence of the gaps, must describe them as well as possible, and must construct the links that plausibly bridge the gaps between the project and the observed impact. The statement of

objectives reveals the impact evaluators' motivation for conducting the study. Did they want to compare different project approaches (learning)? To justify a specific development investment (legitimization/approving accountability)? Did they need information for making management decisions within a current intervention (steering)? By providing the study objectives, the researchers both give the readers a litmus test for the findings (to what degree might they be self-serving) and the measuring stick for the success of the study (did it achieve what it set out to do?).

The IITA study objectives are clearly stated on the bottom of page 4:

> Despite several years of soybean research and diffusion in Nigeria, there remains a dearth of empirical information on the level and extent of its adoption and utilization by resource-poor farmers. No systematic study has investigated the actual impact of soybean on farmers' households and rural communities. The objective of this paper is to assess the real impact of soybean on the lives of small-scale farmers in Benue State, Nigeria. In particular, the paper examines the different gender roles in soybean production, the levels and factors influencing farmers' adoption of improved soybean varieties and utilization innovations, the impact of soybean production on farmers' income generation and distribution, their use and allocation of resources, their household food consumption, nutritional status, welfare, and other relevant social processes in the community.

The main goal of the study, then, is to increase knowledge about the benefits and impact of soybean production for small-scale, resource-poor farmers in selected communities.

However, the IITA study does not explicitly state its limitations. We propose that the main limitation of the study in terms of establishing plausibility is that it does not openly address the attribution gap. (Actually, our workgroup finds two main attribution gaps in the study. Referring to our understanding of the study's operative model of impact, the gaps would lie between "adoption" and "increased income and food security", as well as between "increased income and food security" and "improved well-being of farmers and communities".)

4. The theory of action on which the intervention or strategy has been based.
Development interventions do not perform purely random activities, but their activities are carried out for a purpose. Even when the theory of action has not been made explicit in project documents, the activities are carried out because it is believed that they support desired changes that lead to achieving development goals. Put another way, the source or generator of impact has carried out activities in order to achieve certain outcomes and

impact. The theory of action is the source's construct for how its activities relate to desired impact.

It is the impact evaluator's job to make explicit the reasons the intervention was introduced – what its designers and implementers hope(d) to achieve with the intervention and in what way the intervention activities are/were to contribute to reaching its objectives. Sometimes a stated theory of action has not been adhered to. It then is the duty of the impact evaluator to uncover and present the operative theory of change.

If the impact observed by the study was foreseen by the theory of action, this supports the establishment of plausibility. If not, then the researchers need to clearly address this in their argumentation to establish the plausibility of the linkages between the intervention and the impact actually observed. They then need to answer the question satisfactorily of why the measure or strategy led to unanticipated impact or did not lead to the impact that was foreseen.

The IITA study does not make the theory of action for the observed impact clear. If the intention behind introducing improved soybean technology in Nigeria was ultimately to meet certain (in the report unspecified) socioeconomic objectives as well as to satisfy the needs and priorities of the target population (p. 6), how were each of the related activities (research on soybean varieties, extension services efforts in teaching farmers about new production technologies, etc.) tailored to meet these ends? In the minds of IITA researchers and policy makers, or of extension services agents, how were their activities linked to immediate and/or further-reaching development goals? If we were told, for instance, that extension workers consciously targeted women farmers for receiving seeds and training, this would make the argument more plausible that it was the new soybean varieties and technologies that increased women's socioeconomic status in the region. The reasoning here is that if well-thought-out activities are directed towards achieving certain changes, which then actually do occur, then it is likely that the activities have played a role in bringing about the changes. Of course, this requires clear identification of the source or generator of the impact being examined.

5. The impact hypotheses that the study tests. Whereas the theory of action is an intrinsic part of the intervention or strategy, the impact hypotheses are constructed by the evaluator. Impact hypotheses are statements about the impact that the researcher expects to find. In

other words, the theory of action drives the intervention, and the impact hypotheses drive the assessment. The hypotheses might overlap with the theory of action completely, somewhat, or not at all. The study, then, tests the impact hypotheses.

The GTZ impact model presented earlier can help to illustrate the role played by impact hypotheses. For instance, starting at "use of output" and moving up the impact continuum, impact hypotheses are statements about how the evaluator believes that "use of output" relates to "direct benefits of project", how these in turn relate to or generate "possible indirect benefit of the project", and how the possible indirect benefits feed into the "highly aggregated development progress". Thus impact hypotheses are the constructs with which the evaluators or researchers first attempt to bridge the attribution gap. Data is then collected to check whether the hypotheses plausibly hold along the impact continuum or not.

The IITA study does not present its hypotheses explicitly, or at least not in detail. The main hypothesis is that the soybean sector – specifically the adoption of soybean production and processing technology and of improved varieties – has had positive impact on household income generation and distribution, material welfare, human capital development, children's nutritional status, gender relations, resource use, and social equity within the community (p. 6). But the study presents lists of impact areas and indicators to the exclusion of well-articulated impact hypotheses. If the researchers had developed and pursued their impact hypotheses more explicitly, the paper would have gained in establishing plausibility.

6. Other factors that could have affected the observed changes and alternative impact hypotheses. The attribution gap concept underlines the fact that many factors are at play in effecting the changes observed at a highly-aggregated level – that the changes cannot be reliably and proportionately traced back to any one intervention or program. Arguing believably or plausibly for the linkage between an intervention and observed impact requires that other, significant, influencing factors be addressed and weighed in the balance. Of course, not all other influences can be treated, as there are so many. But the impact evaluator should identify and discuss at least the main factors other than the program or strategy that also could have affected the observed changes and impacts.

Similarly, the researcher should develop alternative hypotheses and demonstrate why the one(s) he or she supports are the most plausible.

Plausibility is then established when the researcher can make a good case for the relationship he or she posits between the intervention and the impact, and why this relationship is meaningful in the context of other major factors involved, and of competing hypotheses. The principals at work here are maximal probability of causality and maximal explanatory value. That is, in comparison to alternative factors, is the study's proposed source of impact likely to have generated the observed impact to some degree? And do the proposed impact hypotheses sufficiently explain the impact relationship in comparison to competing hypotheses?

Obviously, the introduction of soybean and related technologies to Benue State was not the only change experienced by farmers in the region. The IITA study, for instance, only briefly mentions that some experts believed the effects of new policy to have caused much of the impact that was observed (p. 23). The study also states that being closer to a marketplace is associated with higher adoption rates – could the proximity of the market also explain some of the impacts? If the IITA study had explicitly considered the important alternative factors and impact hypotheses, and had defended its own findings in juxtaposition to them, it would have gone a long way towards establishing the credibility and plausibility of the impact it found.

7. Other informed opinions that support and contest the study findings. Plausibility is built upon the expert, informed opinion of the impact evaluator. But he or she is not the only expert in the field. By outlining how other informed people agree with the evaluator's argumentation, the plausibility of the impact relationships is strengthened. Especially important here are the views of the beneficiaries and target groups that support the impact assessment findings, as they are the people who have been directly affected.

The IITA study includes several positive statements of the beneficiaries (pp. 16-18, 26, and 27), that greatly increase the understanding of how soybean farming affected their lives and well-being. The statements also supported the plausible relationship between the introduction of improved varieties/new technology and the socioeconomic impact observed. The study also is sprinkled throughout with references to the research of others that corroborates its findings.

But even the best of studies has its detractors, and even the best of interventions has its "losers", or those who in some way have been disadvantaged by the changes it has

introduced. Given this, plausibility is strengthened when these dissenting viewpoints are presented and countered. How and why do others disagree with the findings of the study? Why are the study's statements about impact nonetheless plausible?

The IITA study is nearly silent as to the opinions of the experts than run counter to its findings. The few exceptions are brief references to varying views on the proportion of farm work performed by men and women (p. 12), the degree of integration of soybean into the local diet (p. 20), the role of policy-shifts versus new technology in the expansion of the soybean sector (p. 23), and whether the expansion of cash crops negatively impact the production of staple food (p. 24). However, what we missed most were the dissenting voices and statements from the "losers" in the study communities, especially to put into context the positive quotes of soybean farmers mentioned above. Did any of the community members disagree or experience negative impacts? If so, why should we believe that the positive impact reported is significant? If not, wonderful! – but we need to be told.

LESSONS FROM THE IITA STUDY AND CONCLUSIONS

The IITA study takes a broader perspective on impact than do most impact assessment reports. It examines a fairly long chain of impact: from the introduction of new soybean varieties and technologies, through diffusion and adoption of the innovations, on to a broad range of socioeconomic effects at the household, farm, and community levels. We can only laud the study's deep and broad view towards impact.

Nevertheless, we find that the study often misses opportunities to construct a sound argument for the plausibility of the relationships posited between the introduction of soybean innovations and the impact observed in the study communities. The two main shortcomings are:

1. The study is not transparent about its central limitation: that attribution gaps – or areas where many other factors affect the observed areas of impact – exist amongst the impact linkages.
2. Taking on the task of *proving* the impact of new soybean varieties and technologies has led to the over-interpretation of data. Information on the highly-aggregated changes in the soybean sector, in the socioeconomic well-being of soybean-farming households, in the allocation of resources, and in the attitudes, values and beliefs prevailing in the study communities is used as proof of the impact of the soybean innovations.

The IITA study forces its data into the role of proving the impact of just one factor amongst the many that generated the changes, thus exaggerating the effects of the new soybean varieties and technologies. This not only undermines credibility, but also misses opportunities for learning. In contrast, the study would have won on credibility and have been able to draw out lessons that would be transferable to other interventions if it had defined its task as one of developing plausible links between the soybean innovations and the changes observed in the study region.

We are in no way implying that the findings of the IITA study are false. Nor do we want to give the impression that lack of attention to what we call the attribution gap is unique to this IITA study. On the contrary: the large majority of impact studies seem to see their task to be the precise measurement and proof of the impact of development efforts. This forces the studies to gloss over the reality of the situation, that the impact on the ground of a given project, program, or strategy cannot be precisely known. By accepting that the best most assessments can do is to establish plausible linkages to impact, they then do not need to gloss over the facts, and their findings therefore become more accurate and useful.

The main message of this paper is that efforts to assess the impact of development work must accept and come to terms with the fact that impact effects cannot be reliably and completely attributed to any one cause. The best that an impact assessment can do is to establish plausible linkages between the project, program or strategy under investigation, and the impact observed on the ground.

This paper proposes a number of standards for establishing plausibility in impact assessment. It turns out that they are not truly new, but relate to the standards for evaluation in general and for the quality of conclusions in qualitative research. They are:

1. Identification of the source of the impact being investigated
2. Presentation of the model or concept of impact used by the impact evaluators and how it applies to the case at hand
3. Statement of the objectives and limitations of the impact assessment
4. Outline of the specific theory of action on which the intervention or strategy has been based
5. Statement of the impact hypotheses that the impact assessment tests
6. Presentation of other factors that could have affected the observed changes and alternative impact hypotheses

7. Discussion of other informed opinions that support and contest the study findings.

We hope that these proposed standards are a pragmatic help to impact evaluators, and that the standards will be refined and expanded through further impact assessment practice.

Contact:
Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH
Dr. Stephan Krall
P.O.Box 5180
65726 Eschborn
Tel. +49 6196 79 1416
Fax +49 6196 79 7137
stephan.krall@gtz.de

Attachment 1: "Standards for Evaluation"

(Summarized in English from: Deutsche Gesellschaft für Evaluation. 2001. "Standards für Evaluation". Anlage 1 zur Beschlussvorlage des Vorstands, Mitgliederversammlung 2001.)

Utility

The utility standards should ensure that the evaluation fulfills the purposes and needs of those for whom the study is intended.

- Identification of the stakeholders.
- Clarification of the purpose of the evaluation.
- Evaluator credibility and competence.
- Appropriate selection and scope of information – broad enough to address the evaluation questions and to be responsive to the information needs of the client and other specified stakeholders.
- Transparency of values.
- Completeness and clarity of the evaluation report.
- Timeliness of the evaluation.
- Evaluation design that maximizes its utilization and impact.

Feasibility

The feasibility standards should ensure that the evaluation is realistic, prudent, diplomatic, and frugal.

- Appropriateness of evaluation process -- correlation between the disruption caused to evaluation participants and the expected value of its findings for them.
- Diplomatic structuring of the evaluation process to maximize later acceptance of the findings.
- Cost effectiveness.

Fairness

The fairness standards should ensure that the evaluation participants are treated fairly and with respect.

- Formal agreement on mutual obligations between the evaluation commissioner and the evaluator.
- Protection of individual rights of study participants.

- Complete and fair assessment.
- Impartial execution of the evaluation and presentation of the findings.
- Accessibility of the findings to the study participants.

Accuracy

The standards on accuracy should ensure that the evaluation provides adequate and valid information.

- Description of the object of the evaluation.
- Context analysis.
- Description of evaluation purposes and procedures.
- Identification/description of information sources.
- Valid and reliable information.
- Systematic checking of findings for mistakes.
- Analysis of quantitative and qualitative information.
- Justification of the conclusions.
- Appropriate documentation the evaluation and archiving of the report.

Attachment 2: "Standards for the Quality of Conclusion"

(Summarized from: Matthew B. Miles and A. Michael Huberman. 1994. Qualitative Data Analysis. 2nd Edition. Thousand Oaks: Sage Publications. Pp. 278-80.)

Objectivity/Confirmability

The main issue is the relative neutrality of the study findings.

- Are the conclusions free from unstated bias?
- Are the study's general methods and procedures described explicitly and in detail?
- Are the conclusions explicitly linked with the data?
- Were competing hypotheses or rival conclusions really considered? Do other rival conclusions seem plausible?
- Are study data retained and available for reanalysis by others?

Reliability/Dependability/Auditability

The underlying concern is whether the study's process is consistent and reasonably stable over time, across researchers, and across methods.

- Have things been done with care?
- Are the research questions clear? Are the features of the study design congruent with the research questions?
- Do findings show meaningful parallelism across data sources?
- Are basic paradigms and analytic constructs clearly specified? (Connection to theory.)
- Were data collected across the full range of appropriate settings, times, respondents, and so on suggested by the research questions?
- Were data quality checks made (e.g., for bias, deceit, informant knowledgeability)?
- Were any forms of peer or colleague review in place?

Internal Validity/Credibility/Authenticity

The question here regards truth-value of the findings.

- Do the findings of the study make sense?
- Are they credible to the people studied? To the readers?
- Does the account of the information leading to the conclusions "ring true"?
- Did triangulation among complementary methods and data source produce generally converging conclusions? if not, is there a coherent explanation for this?

- Are the findings internally coherent?
- Are the concepts systematically related?
- Are areas of uncertainty identified?
- Was negative evidence looked for? Was it found? What happened then?
- Were any predictions made in the study? If so, how accurate were they?

External Validity/Transferability/Fittingness

The issue here is whether the findings have larger import beyond the study case(s).

- Are the conclusion of the study transferable to other contexts?
- Are the characteristics of the original sample, settings, and processes fully described enough to permit adequate comparisons with others?
- Does the report examine possible threats to generalizability?
- Does the researcher define the scope and the boundaries of reasonable generalization from the study?
- Have others reported the findings to be consistent with their own experience? Have the findings been replicated in other studies?
- How do the findings relate to prior theory?
- Is the transferable theory behind the study made explicit?
- Does the report suggest settings where the findings could be tested further?

Utilization/Application/Action Orientation

Apart from the question of transferability of the study findings is the issue of whether the study was an enriching experience for its participants.

- What does the study do for its researchers? For those who were studied?
- Are the findings intellectually and physically accessible to potential users?
- Do the findings stimulate "working hypotheses" on the part of the reader as guidance for future action?
- How can the conclusions be used? (Consciousness raising, empowerment, policy advice, recommendations, etc.)
- Do the findings have a catalyzing effect leading to specific actions?
- Do the actions taken actually help solve the local problem?
- Have the users of the findings learned or developed new capacities?
- Are value-based or ethical concerns raised explicitly in the report? Do some exist implicitly, but that are not addressed by the researcher?